

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 14, 2005, 14:42:49 ; Search time 9009 Seconds
(without alignments)
11036.752 Million cell updates/sec

Title: US-09-806-197-1
Perfect score: 2052
Sequence: 1 catgtgtgcctacagcaca.....gcataacattcttaaatc 2052

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*
2: gb_hcg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|------------|--------------------|
| 1 | 2052 | 100.0 | 2052 | 6 E40087 | E40087 Plant promo |
| 2 | 2048.8 | 99.8 | 2052 | 6 E40093 | E40093 Plant promo |
| 3 | 2038 | 99.3 | 2056 | 6 E40091 | E40091 Plant promo |
| 4 | 2034 | 99.1 | 2048 | 6 E40090 | E40090 Plant promo |
| 5 | 2032.4 | 99.0 | 2048 | 6 E40089 | E40089 Plant promo |
| 6 | 241.2 | 11.8 | 2042 | 6 AR076817 | AR076817 Sequence |
| 7 | 241.2 | 11.8 | 2042 | 6 E15125 | E15125 Promoter. 7 |
| 8 | 196.6 | 9.6 | 247 | 6 AR076816 | AR076816 Sequence |
| 9 | 196.6 | 9.6 | 247 | 6 E15124 | E15124 Promoter. 7 |
| 10 | 195.6 | 9.5 | 246 | 6 AR146852 | AR146852 Sequence |
| 11 | 195.6 | 9.5 | 246 | 6 E55065 | E55065 Plant promo |
| 12 | 174 | 8.5 | 8005 | 8 DCA18706 | Y18706 Daucus caro |
| 13 | 162.4 | 7.9 | 2831 | 6 BD188672 | BD188672 Promotor |
| 14 | 162.4 | 7.9 | 2865 | 6 BD188674 | BD188674 Promotor |
| 15 | 162.4 | 7.9 | 4886 | 8 DARGCHS2 | D1625 Carrot qCHS |
| 16 | 119 | 5.8 | 140 | 6 AR146856 | AR146856 Sequence |
| 17 | 119 | 5.8 | 140 | 6 AR146857 | AR146857 Sequence |
| 18 | 119 | 5.8 | 140 | 6 E55069 | E55069 Plant promo |
| 19 | 119 | 5.8 | 140 | 6 E55070 | E55070 Plant promo |

| | | | | | | |
|------|-------|-----|--------|---|-----------|--------------------|
| c 20 | 106.6 | 5.2 | 250029 | 3 | AE014839 | Plasmodi |
| c 21 | 103.6 | 5.0 | 172816 | 9 | AC093899 | Homo sapi |
| c 22 | 101.8 | 5.0 | 258658 | 3 | AE014832 | Plasmodi |
| c 23 | 100.4 | 4.9 | 348174 | 3 | CR382399 | Plasmodi |
| c 24 | 99.4 | 4.8 | 8056 | 6 | AX599046 | Sequence |
| c 25 | 99.2 | 4.8 | 205130 | 2 | AC105425 | Homo sapi |
| c 26 | 97 | 4.7 | 14867 | 3 | AE001398 | Plasmodi |
| c 27 | 96.6 | 4.7 | 67970 | 3 | PFMAL1P3 | Plasmodi |
| c 28 | 96.4 | 4.7 | 4601 | 3 | DMU11584 | U1584 Drosophila |
| c 29 | 96.4 | 4.7 | 19517 | 3 | DMU37541 | U37541 Drosophila |
| c 30 | 95.8 | 4.7 | 64394 | 9 | AL928596 | Human DNA |
| c 31 | 95.2 | 4.6 | 313050 | 3 | PF3929352 | Plasmodi |
| c 32 | 93.8 | 4.6 | 143331 | 9 | AC091214 | Homo sapi |
| c 33 | 93 | 4.5 | 349980 | 6 | AX344555 | Sequence |
| c 34 | 92.6 | 4.5 | 180629 | 2 | CR388025 | Danio rer |
| c 35 | 92.6 | 4.5 | 349751 | 3 | PFMAL4P3 | Plasmodi |
| c 36 | 92.4 | 4.5 | 261771 | 2 | CR356223 | Danio rer |
| c 37 | 91.8 | 4.5 | 8056 | 6 | AX598900 | Sequence |
| c 38 | 91.6 | 4.5 | 8056 | 6 | AX599046 | Sequence |
| c 39 | 91 | 4.4 | 21153 | 8 | YSCWTCG13 | L36897 Saccharomyc |
| c 40 | 91 | 4.4 | 118642 | 9 | AC126283 | Homo sapi |
| c 41 | 91 | 4.4 | 178427 | 2 | CR392006 | Danio rer |
| c 42 | 90.8 | 4.4 | 86826 | 3 | PFMAL3P5 | Plasmodi |
| c 43 | 90.6 | 4.4 | 155106 | 9 | AC104063 | Homo sapi |
| c 44 | 90.6 | 4.4 | 178670 | 9 | AC104073 | Homo sapi |
| c 45 | 90.6 | 4.4 | 206038 | 2 | CR847802 | Danio rer |

ALIGNMENTS

| RESULT 1 | E40087 | 2052 bp | DNA | linear | PAT 31-JAN-2002 |
|------------|---|--------------------------------|-----|--------|-----------------|
| LOCUS | E40087 | Plant promoter and terminator. | | | |
| DEFINITION | E40087 | | | | |
| ACCESSION | E40087.1 | GI:18627203 | | | |
| VERSION | JP 2000166577-A/1. | | | | |
| KEYWORDS | unidentified | | | | |
| SOURCE | unclassified | | | | |
| ORGANISM | unclassified | | | | |
| REFERENCE | 1 (bases 1 to 2052) | | | | |
| AUTHORS | Nishikawa, S. and Oeda, K. | | | | |
| TITLE | Plant promoter and terminator | | | | |
| JOURNAL | Patent: JP 2000166577-A 1 20-JUN-2000; | | | | |
| COMMENT | SUMITOMO CHEM CO LTD | | | | |
| | OS Daucus carota L. | | | | |
| | PN JP 2000166577-A/1 | | | | |
| | PD 20-JUN-2000 | | | | |
| | PF 01-OCT-1999 | | | | |
| | PR JP 1999281475 | | | | |
| | PI SATOMI NISHIKAWA, KENJI OEDA | | | | |
| | PC C12N15/09,A01H5/00,C12N1/21,C12N5/10/(C12N5/10,C12R1.91), PC | | | | |
| | C12N15/00, | | | | |
| | PC C12N5/00, (C12N5/00,C12R1.91) | | | | |
| | CC | | | | |
| | PH Key | | | | |
| | FT promoter | | | | |
| | Location/Qualifiers | | | | |
| | 1..2052 | | | | |
| | /organism="unidentified" | | | | |
| | /mol_type="genomic DNA" | | | | |
| | /db_xref="taxon:32644" | | | | |

Query Match 100.0%; Score 2052; DB 6; Length 2052;
Best Local Similarity 100.0%; Pred. No. 4.4e-310;
Matches 2052; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CATGTGTGCCTACACACATAGGCTGTTCGTTGAGAGACGACAGCTGCTCTGA 60
|||||
Db 1 CATGTGTGCCTACACACATAGGCTGTTCGTTGAGAGACGACAGCTGCTCTGA 60
|||||

| | | | |
|----|------|--|------|
| Qy | 61 | CTTCTTCTTCTTTGACCTGTTTGTATAAAGAGCTAGAAATATTTTTTAAAGAGCTGCGAA | 120 |
| Db | 61 | CTTCTTCTTCTTTTGACCTGTTTGTATAAAGAGCTAGAAATATTTTTTAAAGAGCTGCGAA | 120 |
| Qy | 121 | TACTAACTTCTCTCTCAACAACCTCCGGCTCTTTTCCAAACACCTTTATTTAACTTTTCTT | 180 |
| Db | 121 | TACTAACTTCTCTCTCAACAACCTCCGGCTCTTTTCCAAACACCTTTATTTAACTTTTCTT | 180 |
| Qy | 181 | TCTCATTTTCTACTCCACTTCTTCTCTATAAGCAAGAAATCACCTCTTTTAAAGCTAACCCA | 240 |
| Db | 181 | TCTCATTTTCTACTCCACTTCTTCTCTATAAGCAAGAAATCACCTCTTTTAAAGCTAACCCA | 240 |
| Qy | 241 | AACGGCTCAATAAAGATCATTCATTAATGTATCTTTCAATTTTAGGATAACAATACGT | 300 |
| Db | 241 | AACGGCTCAATAAAGATCATTCATTAATGTATCTTTCAATTTTAGGATAACAATACGT | 300 |
| Qy | 301 | GAAACAGGGTTATTTTTTAAACGGTGCAACAATCTCTAATAATTTTACCTGCCGGTGAACA | 360 |
| Db | 301 | GAAACAGGGTTATTTTTTAAACGGTGCAACAATCTCTAATAATTTTACCTGCCGGTGAACA | 360 |
| Qy | 361 | CCGTCTTCCAGATAATATATTTTAAATTTTGTAGCCTCCCTTTTAAACCAATTCGCATGC | 420 |
| Db | 361 | CCGTCTTCCAGATAATATATTTTAAATTTTGTAGCCTCCCTTTTAAACCAATTCGCATGC | 420 |
| Qy | 421 | AGGACGACTTAGGTGAATACACATGTGTAGTCTTTTAAACAAAGAAACAAGTGGTTC | 480 |
| Db | 421 | AGGACGACTTAGGTGAATACACATGTGTAGTCTTTTAAACAAAGAAACAAGTGGTTC | 480 |
| Qy | 481 | ATGCTCAGCCATCAAAATTGACAAAACCCGACACAACTCTATCCACGCTACTATCTTT | 540 |
| Db | 481 | ATGCTCAGCCATCAAAATTGACAAAACCCGACACAACTCTATCCACGCTACTATCTTT | 540 |
| Qy | 541 | TGGCCGAATGCTTCTCAAAATGTTTTTATATGTAAATTAATGCCCATCCAAAGGATAAGT | 600 |
| Db | 541 | TGGCCGAATGCTTCTCAAAATGTTTTTATATGTAAATTAATGCCCATCCAAAGGATAAGT | 600 |
| Qy | 601 | AAAAATCCCGTTTAAACGAGTTTGTAAATATATATATATATATATATATATATATATAT | 660 |
| Db | 601 | AAAAATCCCGTTTAAACGAGTTTGTAAATATATATATATATATATATATATATATATAT | 660 |
| Qy | 661 | AATACTTTTAGACAGACAAGAGACTTAGGTCAAAAATGGACGCTGGTAAACAGCCTAGACT | 720 |
| Db | 661 | AATACTTTTAGACAGACAAGAGACTTAGGTCAAAAATGGACGCTGGTAAACAGCCTAGACT | 720 |
| Qy | 721 | TGGTCACGTAAATAGATAATTTGTAGTATAATATAGTAGGATCTACAATGACATTTAA | 780 |
| Db | 721 | TGGTCACGTAAATAGATAATTTGTAGTATAATATAGTAGGATCTACAATGACATTTAA | 780 |
| Qy | 781 | ATTAGAGCTATTAATTAAGTTTACTAATAAATAGAGAGGTTAGTAAACAGAAAGCAGGTA | 840 |
| Db | 781 | ATTAGAGCTATTAATTAAGTTTACTAATAAATAGAGAGGTTAGTAAACAGAAAGCAGGTA | 840 |
| Qy | 841 | AAAAACAGAGCTTGCTGCTGTGTGTAGTTGTGTAGTCAATTTCTTTTAAAGTAATG | 900 |
| Db | 841 | AAAAACAGAGCTTGCTGCTGTGTGTAGTTGTGTAGTCAATTTCTTTTAAAGTAATG | 900 |
| Qy | 901 | TAAACTGATCTAAGACACATAGAAATTTAGTACAGGTTTAAACCTTTTACAGAAATTTATA | 960 |
| Db | 901 | TAAACTGATCTAAGACACATAGAAATTTAGTACAGGTTTAAACCTTTTACAGAAATTTATA | 960 |
| Qy | 961 | TTAAACGAAATCATTTTTATTAACATGTCTCTCGGCTGTCATTATAATAGGGATCACTTAC | 1020 |
| Db | 961 | TTAAACGAAATCATTTTTATTAACATGTCTCTCGGCTGTCATTATAATAGGGATCACTTAC | 1020 |
| Qy | 1021 | TGATCATCTTAATAAACCTTGTTTAAACAAATTCGAATAGGATAAATATCTTACAATGAA | 1080 |
| Db | 1021 | TGATCATCTTAATAAACCTTGTTTAAACAAATTCGAATAGGATAAATATCTTACAATGAA | 1080 |
| Qy | 1081 | AAGAGGACAATGCTCTTTTGAACAAAACAAATAGTACTCCCTCCGCTCCCTCTCAAAATGT | 1140 |
| Db | 1081 | AAGAGGACAATGCTCTTTTGAACAAAACAAATAGTACTCCCTCCGCTCCCTCTCAAAATGT | 1140 |
| Qy | 1141 | ATACATATGGATTTGGACACGGAGACTTAGAAAAATGTATAAGTAATGTAGAGTAAAAAG | 1200 |

[illegible]

RESULT 2:

| | | | | | |
|------------|--------------------------------|-------------|-----|--------|-----------------|
| E40093 | E40093 | 2052 bp | DNA | linear | PAT 31-JAN-2002 |
| LOCUS | | | | | |
| DEFINITION | Plant promoter and terminator. | | | | |
| ACCESSION | E40093 | | | | |
| VERSION | E40093.1 | GI:18627209 | | | |
| KEYWORDS | JP 2000166577-A/7. | | | | |
| SOURCE | unidentified | | | | |
| ORGANISM | unidentified | | | | |

[illegible]

| | | | |
|---|------|--|------|
| Db | 1741 | GTTTGAACAATGATGTCGGGTGACATCTATGACCTTTCAACTCAAACTAGTGAATAAT | 1800 |
| Qy | 1801 | GCATTTCTAGAAATACATCTTTTCAAAATTTCAACAAACACAGCTTTAACTTTCTTTCAACG | 1860 |
| Db | 1801 | GCATTTCTAGAAATACATCTTTTCAAAATTTCAACAAACACAGCTTTAACTTTCTTTCAACG | 1860 |
| Qy | 1861 | GATTGGAATCCTTTTCTTAAACTTTTAAATAAATAAATAATGCAATTAATGTAATATTATC | 1920 |
| Db | 1861 | GATTGGAATCCTTTTCTTAAACTTTTAAATAAATAAATAATGCAATTAATGTAATATTATC | 1920 |
| Qy | 1921 | AACACCTCAACATGATGTTAGGCTACTATAAATAGTGCTCTTGCTGCTCTACTATCAT | 1980 |
| Db | 1921 | AACACCTCAACATGATGTTAGGCTACTATAAATAGTGCTCTTGCTGCTCTACTATCAT | 1980 |
| Qy | 1981 | CACATCAATCTTACACACAAACCTTGAGCTTAATTTTCTACTATTCTCAGCAATAC | 2040 |
| Db | 1981 | CACATCAATCTTACACACAAACCTTGAGCTTAATTTTCTACTATTCTCAGCAATAC | 2040 |
| Qy | 2041 | ATTCTAAATATC 2052 | |
| Db | 2041 | ATTCTAAAGATC 2052 | |
| RESULT 3 | | | |
| LOCUS E40091 2056 bp DNA linear PAT 31-JAN-2002 | | | |
| DEFINITION Plant promoter and terminator. | | | |
| ACCESSION E40091 | | | |
| VERSION E40091.1 GI:18627207 | | | |
| KEYWORDS JP 2000166577-A/5. | | | |
| SOURCE unidentified | | | |
| ORGANISM unidentified | | | |
| REFERENCE 1. (bases 1 to 2056) | | | |
| AUTHORS Nishikawa,S. and Oeda,K. | | | |
| TITLE Plant promoter and terminator | | | |
| JOURNAL Patent: JP 2000166577-A 5 20-JUN-2000; | | | |
| SDMITOMO CHEM CO LTD | | | |
| COMMENT OS Daucus carota L. | | | |
| PN JP 2000166577-A/5 | | | |
| PD 20-JUN-2000 | | | |
| PF 01-OCT-1999 JP 1999281475 | | | |
| PR | | | |
| PI SATOMI NISHIKAWA,KENJI OEDA | | | |
| PC C12N15/09,A01H5/00,C12N1/21,C12N5/10// (C12N5/10,C12R1:91), PC | | | |
| C12N15/00, | | | |
| PC C12N5/00, (C12N5/00,C12R1:91) | | | |
| CC | | | |
| Key Location/Qualifiers | | | |
| FT promoter (1)-(2056). | | | |
| source 1..2056 | | | |
| /organism="unidentified" | | | |
| /mol_type="genomic DNA" | | | |
| /db_xref="taxon:32644" | | | |
| ORIGIN | | | |
| Query Match 99.3%; Score 2038; DB 6; Length 2056; | | | |
| Best Local Similarity 99.8%; Pred. No. 6,7e-308; | | | |
| Matches 2052; Conservative 0; Mismatches 0; Indels 4; Gaps 1; | | | |
| Qy | 1 | CATGTGCGCCTACAGCACATAGGCGCTGTTGGTTGAGAGAGCAGAGCTGCTCTGA | 60 |
| Db | 1 | CATGTGCGCCTACAGCACATAGGCGCTGTTGGTTGAGAGAGCAGAGCTGCTCTGA | 60 |
| Qy | 61 | CTTCTCTCTTTTGACCTGTTGTATAAAGAGTAGAATAATTTTAAAGCTCGAA | 120 |
| Db | 61 | CTTCTCTCTTTTGACCTGTTGTATAAAGAGTAGAATAATTTTAAAGCTCGAA | 120 |
| Qy | 121 | TACTAACTCTCTCAACAATCCGCTTCTTTTCCAAACACTTTATTAATCTTTTACT | 180 |
| Db | 121 | TACTAACTCTCTCAACAATCCGCTTCTTTTCCAAACACTTTATTAATCTTTTACT | 180 |

| | | | |
|----|------|--|------|
| Qy | 181 | TCTCATTTCTACTCCACTTCTTTGCTATAAGCAAGAAATCACTTCTTTTAAGCTAACCCA | 240 |
| Db | 181 | TCTCATTTCTACTCCACTTCTTTGCTATAAGCAAGAAATCACTTCTTTTAAGCTAACCCA | 240 |
| Qy | 241 | AACGGCCTCAATAAAGATCATTTCAATAATGATCTTTCAATTTTAGGATAACAATACGT | 300 |
| Db | 241 | AACGGCCTCAATAAAGATCATTTCAATAATGATCTTTCAATTTTAGGATAACAATACGT | 300 |
| Qy | 301 | GAACAGGGTATTTTAAACGGTCAACAAATTTCTAATAATTTTACCTGGCGGTGAACA | 360 |
| Db | 301 | GAACAGGGTATTTTAAACGGTCAACAAATTTCTAATAATTTTACCTGGCGGTGAACA | 360 |
| Qy | 361 | CGGTCTTCCAAAGATAATATATTTTAAATTTGTAGCTCCCTTTTAAACCAAAATTCGATGC | 420 |
| Db | 361 | CGGTCTTCCAAAGATAATATATTTTAAATTTGTAGCTCCCTTTTAAACCAAAATTCGATGC | 420 |
| Qy | 421 | AGGACGACTTAGGTGAATACACATTTGCTGAGTCTTTTAAACAAAGAAACAGTGGTTC | 480 |
| Db | 421 | AGGACGACTTAGGTGAATACACATTTGCTGAGTCTTTTAAACAAAGAAACAGTGGTTC | 480 |
| Qy | 481 | ATGCTCAGCCATCAAAATTTGACAAACCCGACACAACTCTATCCACGTACTATACTTT | 540 |
| Db | 481 | ATGCTCAGCCATCAAAATTTGACAAACCCGACACAACTCTATCCACGTACTATACTTT | 540 |
| Qy | 541 | TGGCCGAATGCTTCTCAAAATGTTTTTATATGTAATAATAATGCCCATCCAAAGGATAAGT | 600 |
| Db | 541 | TGGCCGAATGCTTCTCAAAATGTTTTTATATGTAATAATAATGCCCATCCAAAGGATAAGT | 600 |
| Qy | 601 | AAAAATCCCGTTAAACCGAGTTGTTTAAATATATATGTTTACCTTACAGAGGATATTCGT | 660 |
| Db | 601 | AAAAATCCCGTTAAACCGAGTTGTTTAAATATATATGTTTACCTTACAGAGGATATTCGT | 660 |
| Qy | 661 | AATACTTTTAGACGACAGAGACTTAGGTCAAAATGAGCGCTGGTAAACAGCCTAGACT | 720 |
| Db | 661 | AATACTTTTAGACGACAGAGACTTAGGTCAAAATGAGCGCTGGTAAACAGCCTAGACT | 720 |
| Qy | 721 | TGGTCACTGATAAATAGATAATTTGTTAGTATATAATAGTAGGATCTACAATGACATTAA | 780 |
| Db | 721 | TGGTCACTGATAAATAGATAATTTGTTAGTATATAATAGTAGGATCTACAATGACATTAA | 780 |
| Qy | 781 | ATTAGAGCTATTAAATTAAGTTACTAATAAATAGAGAGGTTAGTAAACAGAGGATA | 840 |
| Db | 781 | ATTAGAGCTATTAAATTAAGTTACTAATAAATAGAGAGGTTAGTAAACAGAGGATA | 840 |
| Qy | 841 | AAAAAAGAGCTTGCTGCTGTTAGTTAGTCTGTTGAGCTCATTTCTTTTAAAGTAATG | 900 |
| Db | 841 | AAAAAAGAGCTTGCTGCTGTTAGTTAGTCTGTTGAGCTCATTTCTTTTAAAGTAATG | 900 |
| Qy | 901 | TAAACTGATCTAAAGCACATAGAAATTTAGTACAGGTTAAAACTTTTACAAGAAATTTATA | 960 |
| Db | 901 | TAAACTGATCTAAAGCACATAGAAATTTAGTACAGGTTAAAACTTTTACAAGAAATTTATA | 960 |
| Qy | 961 | TTAAACGAAAAATCAATTTTATAACATGCTCTCGGCTGTCATTATAATAGGATCACTTAC | 1020 |
| Db | 961 | TTAAACGAAAAATCAATTTTATAACATGCTCTCGGCTGTCATTATAATAGGATCACTTAC | 1020 |
| Qy | 1021 | TGATCATCCATTAAABACCTTGTAAACAAATTTCAATGAGATAAATAATCTTTACAATGAA | 1080 |
| Db | 1021 | TGATCATCCATTAAABACCTTGTAAACAAATTTCAATGAGATAAATAATCTTTACAATGAA | 1080 |
| Qy | 1081 | AAGAAGGACAAATGCTCTCTTTGAAAAAAACAAATAGGTACTCCCTCCGCTCTCTGAAATGT | 1140 |
| Db | 1081 | AAGAAGGACAAATGCTCTCTTTGAAAAAAACAAATAGGTACTCCCTCCGCTCTCTGAAATGT | 1140 |
| Qy | 1141 | ATACATATCGATTGGACACGGAGCTTAAGAAAAATGTAATAAGTAATAGTAGATAAAG | 1200 |
| Db | 1141 | ATACATATCGATTGGACACGGAGCTTAAGAAAAATGTAATAAGTAATAGTAGATAAAG | 1200 |
| Qy | 1201 | AAAGAGAAAGAAAGTGGGTAAAGTAGCGGGACCCACCAATATATATTTGATAGATTAG | 1260 |
| Db | 1201 | AAAGAGAAAGAAAGTGGGTAAAGTAGCGGGACCCACCAATATATATTTGATAGATTAG | 1260 |

| | | | |
|------------|---|--|------|
| Qy | 1261 | AAAAGTAGTTCAAAAGTAGTGGTGGGATGTTTATATATATATAAATTTACTATTTTG | 1320 |
| Db | 1261 | AAAAGTAGTTGAAAAGTAGTGGTGGGATGTTTATATATATAAATTTACTATTTTG | 1320 |
| Qy | 1321 | AGAAAGTTTTGAAAATGTATAGAAATGAGTGGGACATCCATAAAGGAAAGGTATAGAAT | 1380 |
| Db | 1321 | AGAAAGTTTTGAAAATGTATAGAAATGAGTGGGACATCCATAAAGGAAAGGTATAGAAT | 1380 |
| Qy | 1381 | TAAATGGGACAGAGGGAGTAATACCTTTATGATATATAAATTTTTGTTATTTGATTCA | 1440 |
| Db | 1381 | TAAATGGGACAGAGGGAGTAATACCTTTATGATATATAAATTTTTGTTATTTGATTCA | 1440 |
| Qy | 1441 | TAGATTATAAATCTATGTTTATATGATATATAAATTTTTTAAATATATATATTAAT | 1500 |
| Db | 1441 | TAGATTATAAATCTATGTTTATATGATATATAAATTTTTTAAATATATATATTAAT | 1500 |
| Qy | 1501 | CTGATTAGTGCATTACCGCTTTTATAATTTTCAATCTAGATATATGAATAAATCAG | 1560 |
| Db | 1501 | CTGATTAGTGCATTACCGCTTTTATAATTTTCAATCTAGATATATGAATAAATCAG | 1560 |
| Qy | 1561 | TTATCTGAAAAGCAAAATAATCTTTGTAAACAGCGTTCGGTCAAAATGGGAAGTTCAATG | 1620 |
| Db | 1561 | TTATCTGAAAAGCAAAATAATCTTTGTAAACAGCGTTCGGTCAAAATGGGAAGTTCAATG | 1620 |
| Qy | 1621 | TGTAATCAATAGTTTTTAAATAAAGTAAATTTTAAATTTAAATTTGTTATTTGTTTCA | 1680 |
| Db | 1621 | TGTAATCAATAGTTTTTAAATAAAGTAAATTTTAAATTTAAATTTGTTATTTGTTTCA | 1680 |
| Qy | 1681 | AAATTTAAATAAATTTTGGATGGGAGTTCACGGGCATCATTTGAGCAGCATAGACT | 1740 |
| Db | 1681 | AAATTTAAATAAATTTTGGATGGGAGTTCACGGGCATCATTTGAGCAGCATAGACT | 1740 |
| Qy | 1741 | GTTTGAACAATGTATGTCCGGTGACATCTATGACCTTTTCAACTCAAACTAGTGAATAAT | 1800 |
| Db | 1741 | GTTTGAACAATGTATGTCCGGTGACATCTATGACCTTTCAACTCAAACTAGTGAATAAT | 1800 |
| Qy | 1801 | GCATT---CTAGAATCAATCTTTTCAAAATTTCAACAAACACAGCTTTAATCTTTTC | 1856 |
| Db | 1801 | GCATTCTAGCTAGAATACATCTTTTCAAAATTTCAACAAACACAGCTTTAATCTTTTC | 1860 |
| Qy | 1857 | AACGGATTGGATCCCTTTCTAACTTTTAAATAAATAAATAAATGCAATTTGTAATATT | 1916 |
| Db | 1861 | AACGGATTGGAAATCCCTTTCTAACTTTTAAATAAATAAATAAATGCAATTTGTAATATT | 1920 |
| Qy | 1917 | TATCAACACCTCAACATTTGATTTAGGTGACTATAAATAGTGTCTTTGGTGCTCTACTA | 1976 |
| Db | 1921 | TATCAACACCTCAACATTTGATTTAGGTGACTATAAATAGTGTCTTTGGTGCTCTACTA | 1980 |
| Qy | 1977 | TCATCACATCAATCTTACACCACAACTTTGAGCTTAATTTTTCTACTTATTTCTCAGCAA | 2036 |
| Db | 1981 | TCATCACATCAATCTTACACCACAACTTTGAGCTTAATTTTTCTACTTATTTCTCAGCAA | 2040 |
| Qy | 2037 | TAACTTCTAAATATC 2052 | |
| Db | 2041 | TAACTTCTAAATATC 2056 | |
| RESULT 4 | | | |
| E40090 | | | |
| LOCUS | E40090 2048 bp DNA linear PAT 31-JAN-2002 | | |
| DEFINITION | Plant promoter and terminator. | | |
| ACCESSION | E40090 | | |
| VERSION | E40090.1 GI:18627206 | | |
| KEYWORDS | JP 2000166577-A/4. | | |
| SOURCE | unidentified | | |
| ORGANISM | unclassified. | | |
| REFERENCE | 1 (bases 1 to 2048) | | |
| AUTHORS | Nishikawa, S. and Oeda, K. | | |
| TITLE | Plant promoter and terminator | | |
| JOURNAL | Patent: JP 2000166577-A 4 20-JUN-2000; | | |
| COMMENT | SUMITOMO CHEM CO LTD | | |
| | OS Daucus carota L. | | |

| | | | |
|---|--|--|-----|
| PN | JP | 2000166577-A/4 | |
| PD | 20-JUN-2000 | | |
| PF | 01-OCT-1999 | JP 1999281475 | |
| PR | | | |
| PI | SATOMI NISHIKAWA, KENJI OEDA | | |
| PC | C12N15/09, A01H5/00, C12N1/21, C12N5/10// (C12N5/10, C12R1:91), PC | | |
| PC | C12N15/00, (C12N5/00, C12R1:91) | | |
| PC | C12N5/00, (C12N5/00, C12R1:91) | | |
| CC | | | |
| FH | Key | Location/Qualifiers | |
| FT | promoter | (1). (2048). | |
| FEATURES | | | |
| source | | | |
| ORIGIN | | | |
| Query Match 99.1%; Score 2034; DB 6; Length 2048; | | | |
| Best Local Similarity 99.8%; Pred. No. 2.8e-307; | | | |
| Matches 2048; Conservative 0; Mismatches 0; Indels 4; Gaps 1; | | | |
| Qy | 1 | CATGTGTGCCCTACAGCACATAGGCGCTGTTGGTTGGAGAGAGCAGAGCTGCTCTGA | 60 |
| Db | 1 | CATGTGTGCCCTACAGCACATAGGCGCTGTTGGTTGGAGAGAGCAGAGCTGCTCTGA | 60 |
| Qy | 61 | CTTCTCTCTCTTTGACCTGTTGTATAAAGAGTAGAATAATTTTTAAAAAGCTGCGAA | 120 |
| Db | 61 | CTTCTCTCTCTTTGACCTGTTGTATAAAGAGTAGAATAATTTTTAAAAAGCTGCGAA | 120 |
| Qy | 121 | TACTAACTTCTCTCACAACTTCCGCTTCTTTTCCAAACACATTTTAACTTTTACT | 180 |
| Db | 121 | TACTAACTTCTCTCACAACTTCCGCTTCTTTTCCAAACACATTTTAACTTTTACT | 180 |
| Qy | 181 | TCTCATTTCTCTCCACTTCTTGTATAGCAAGAAATCACTTCTTTTAAAGCTAAACCA | 240 |
| Db | 181 | TCTCATTTCTCTCCACTTCTTGTATAGCAAGAAATCACTTCTTTTAAAGCTAAACCA | 240 |
| Qy | 241 | AACGGCTCAATAAAGATCAATTCATAAATGTATCTTTCAATTTTAGGATAACAACTAGT | 300 |
| Db | 241 | AACGGCTCAATAAAGATCAATTCATAAATGTATCTTTCAATTTTAGGATAACAACTAGT | 300 |
| Qy | 301 | GAACAGGGTTATTTTAAACGTGTCAACAAATTTCTAATAATTTTAACTGCGCGGTGAACA | 360 |
| Db | 301 | GAACAGGGTTATTTTAAACGTGTCAACAAATTTCTAATAATTTTAACTGCGCGGTGAACA | 360 |
| Qy | 361 | CGCTCTCCAGATAATATATTTTAAATTTTGTAGCTCCCTTTTAAACAAATTCGCGATGC | 420 |
| Db | 361 | CGCTCTCCAGATAATATATTTTAAATTTTGTAGCTCCCTTTTAAACAAATTCGCGATGC | 420 |
| Qy | 421 | AGGACGACTTAGGTGAATACACATTTGCTGAGTCTTTTAAACAAAGAACAGTGGTTC | 480 |
| Db | 417 | AGGACGACTTAGGTGAATACACATTTGCTGAGTCTTTTAAACAAAGAACAGTGGTTC | 476 |
| Qy | 481 | ATGCTCAGCCATCAAAATTTGACAAACCCGACACACACTCTATCCAGTACTATCTTT | 540 |
| Db | 477 | ATGCTCAGCCATCAAAATTTGACAAACCCGACACACACTCTATCCAGTACTATCTTT | 536 |
| Qy | 541 | TGCGCCGATGCTTCTCAAAATGTTTTTATATGATTAATAATGCCCCATCCAGGATAGT | 600 |
| Db | 537 | TGCGCCGATGCTTCTCAAAATGTTTTTATATGATTAATAATGCCCCATCCAGGATAGT | 596 |
| Qy | 601 | AAATTTCCCGTTTAAACAGTTTGTATATATATATATATATATATATATATATATATAT | 660 |
| Db | 597 | AAATTTCCCGTTTAAACAGTTTGTATATATATATATATATATATATATATATATATAT | 656 |
| Qy | 661 | AATACTTTTAGACGACAAAGAGACTTAGGTCAAAATGAGCGCTGGTAAACAGCGTAGACT | 720 |
| Db | 657 | AATACTTTTAGACGACAAAGAGACTTAGGTCAAAATGAGCGCTGGTAAACAGCGTAGACT | 716 |
| Qy | 721 | TGCTACTGATAAATAGATAATTTGTTAGTATAATATATATATATATATATATATATATAT | 780 |
| Db | 717 | TGCTACTGATAAATAGATAATTTGTTAGTATAATATATATATATATATATATATATATAT | 776 |

| | | | |
|---|------|---|------|
| QY | 781 | ATTAGAGCTAATTAAGTTACTAATAAATAAGAGAGGTTAGTAAACAGAAAGCAGGTA | 840 |
| Db | 777 | ATTAGAGCTAATTAAGTTACTAATAAATAAGAGAGGTTAGTAAACAGAAAGCAGGTA | 836 |
| QY | 841 | AAACAGAGCTTGCTGCTGTGTGTTAGTTGTTGAGCTCAATTTCTTTTAAAGTAATG | 900 |
| Db | 837 | AAACAGAGCTTGCTGCTGTGTGTTAGTTGTTGAGCTCAATTTCTTTTAAAGTAATG | 896 |
| QY | 901 | TAACTGATCTAAGCAGCATAGAAATTTAGTACAGGTTAAACCTTTTACAGAAATTTATA | 960 |
| Db | 897 | TAACTGATCTAAGCAGCATAGAAATTTAGTACAGGTTAAACCTTTTACAGAAATTTATA | 956 |
| QY | 961 | TTAAACGAAATCAATTTTATAACAATGTCTCCGGCTGTCTATTATAATAGGGATCACTTAC | 1020 |
| Db | 957 | TTAAACGAAATCAATTTTATAACAATGTCTCTCGGCTGTCTATTATAATAGGGATCACTTAC | 1016 |
| QY | 1021 | TGATCATCCATTAACCTTTGTTAAACAAATCAATGAGATAAAATATCTTACAATGAA | 1080 |
| Db | 1017 | TGATCATCCATTAACCTTTGTTAAACAAATCAATGAGATAAAATATCTTACAATGAA | 1076 |
| QY | 1081 | AAGAAGCAGATGCTCTTTGAAAAACAATAAGTACTCCCTCCGCTCCTCTGAAATGT | 1140 |
| Db | 1077 | AAGAAGCAGATGCTCTTTGAAAAACAATAAGTACTCCCTCCGCTCCTCTGAAATGT | 1136 |
| QY | 1141 | ATACATATGATTGGACACCGAGACTAAGAAAAATGTATAAAGTAAATGTAGAGTAAAAAG | 1200 |
| Db | 1137 | ATACATATGATTGGACACCGAGACTAAGAAAAATGTATAAAGTAAATGTAGAGTAAAAAG | 1196 |
| QY | 1201 | AAAGAGAAAGAAAGTGGGTAAAGTAGCGGGACCCCAATATATTAATGTAGATGATTAG | 1260 |
| Db | 1197 | AAAGAGAAAGAAAGTGGGTAAAGTAGCGGGACCCCAATATATTAATGTAGATGATTAG | 1256 |
| QY | 1261 | AAAAGTAGTTGAAAGTAGTGGGTGGGTGGGATTTTATATATATAAAATTTTACTATTTTG | 1320 |
| Db | 1257 | AAAAGTAGTTGAAAGTAGTGGGTGGGTGGGATTTTATATATAAAATTTTACTATTTTG | 1316 |
| QY | 1321 | AGAAAGTTTGAATGTATAGAAATGAGTGGGACATCCATAAAGGAAAGTGATAGAAT | 1380 |
| Db | 1317 | AGAAAGTTTGAATGTATAGAAATGAGTGGGACATCCATAAAGGAAAGTGATAGAAT | 1376 |
| QY | 1381 | TAAATGGCAGAGGGAGTAATACCTTTATGATATATAAATTTTGTATTTTGATTCTCA | 1440 |
| Db | 1377 | TAAATGGCAGAGGGAGTAATACCTTTATGATATATAAATTTTGTATTTTGATTCTCA | 1436 |
| QY | 1441 | TAGATTATAAACTATGTTATAATGATAATATAATTTTAAAAATAATCTATATAATT | 1500 |
| Db | 1437 | TAGATTATAAACTATGTTATAATGATAATATAATTTTAAAAATAATCTATATAATT | 1496 |
| QY | 1501 | CTGATTAGTCGATTACCGCCTTTTATAATTTTACAATGAGTAAATATGATAAATATCAG | 1560 |
| Db | 1497 | CTGATTAGTCGATTACCGCCTTTTATAATTTTACAATGAGTAAATATGATAAATATCAG | 1556 |
| QY | 1561 | TTATCTGAAAGCAAAATAATCTTTGTAAAAACAGCGTTCCGGTCAAAATGGGAAGTTCAATG | 1620 |
| Db | 1557 | TTATCTGAAAGCAAAATAATCTTTGTAAAAACAGCGTTCCGGTCAAAATGGGAAGTTCAATG | 1616 |
| QY | 1621 | TGTAATCAATAGTTTTTAATAATAAAGTAAATTTTAAATTAATTTGTTATTTTGTTCAGA | 1680 |
| Db | 1617 | TGTAATCAATAGTTTTTAATAATAAAGTAAATTTTAAATTAATTTGTTATTTTGTTCAGA | 1676 |
| QY | 1681 | AAATTAATAATAATTTATGAGCATGGGAAGTTACCGGCGATCATGTAGCAGCACTAGACT | 1740 |
| Db | 1677 | AAATTAATAATAATTTATGAGCATGGGAAGTTACCGGCGATCATGTAGCAGCACTAGACT | 1736 |
| QY | 1741 | GTTTGAACAATGTATGTCCGGGTGACATCTATGACCTTTTCAACTCAAACTAGTGAATAAT | 1800 |
| Db | 1737 | GTTTGAACAATGTATGTCCGGGTGACATCTATGACCTTTTCAACTCAAACTAGTGAATAAT | 1796 |
| QY | 1801 | GCATTTAGAAATACATCTTTTCAAAATTTTCAACAAAACACAGCTTTTAACTTTTCTTCAACG | 1860 |
| Db | 1797 | GCATTTAGAAATACATCTTTTCAAAATTTTCAACAAAACACAGCTTTTAACTTTTCTTCAACG | 1856 |
| QY | 1861 | GATTGGAATCCTTTTCTTAACTTTTAAATAAAAAAATGCATTATTGTAATATTTATC | 1920 |
| Db | 1857 | GATTGGAATCCTTTTCTTAACTTTTAAATAAAAAAATGCATTATTGTAATATTTATC | 1916 |
| QY | 1921 | AACACCTCAACATTTGATTTAGCTACTATAAATAGGTGCTCTTGGTGTCTTACTATCAT | 1980 |
| Db | 1917 | AACACCTCAACATTTGATTTAGCTACTATAAATAGGTGCTCTTGGTGTCTTACTATCAT | 1976 |
| QY | 1981 | CACATCAATCTTTACACCAACAACTTGAGCTTAAATTTTCTACTATTCTCAGCAATAAC | 2040 |
| Db | 1977 | CACATCAATCTTTACACCAACAACTTGAGCTTAAATTTTCTACTATTCTCAGCAATAAC | 2036 |
| QY | 2041 | ATTCTAAATATC 2052 | |
| Db | 2037 | ATTCTAAATATC 2048 | |
| RESULT 5 | | | |
| E40089 | | | |
| LOCUS | | | |
| DEFINITION | | | |
| E40089 | | | |
| ACCESSION | | | |
| E40089.1 | | | |
| VERSION | | | |
| JP 2000166577-A/3. | | | |
| KEYWORDS | | | |
| unidentified | | | |
| SOURCE | | | |
| unclassified. | | | |
| ORGANISM | | | |
| 1 (bases 1 to 2048) | | | |
| REFERENCE | | | |
| Nishikawa, S. and Oeda, K. | | | |
| AUTHORS | | | |
| Plant promoter and terminator | | | |
| TITLE | | | |
| Patent: JP 2000166577-A 3 20-JUN-2000; | | | |
| JOURNAL | | | |
| SUMITOMO CHEM CO LTD | | | |
| COMMENT | | | |
| OS Daucus carota L. | | | |
| PN JP 2000166577-A/3 | | | |
| PD 20-JUN-2000 | | | |
| PR 01-OCT-1999 JP 1999281475 | | | |
| PI SATOMI NISHIKAWA, KENJI OEDA | | | |
| PC C12N15/09, A01H5/00, C12N1/21, C12N5/10// (C12N5/10, C12R1:91), PC | | | |
| C12N15/00, | | | |
| PC C12N5/00, (C12N5/00, C12R1:91) | | | |
| CC | | | |
| FH Key | | | |
| FT promoter | | | |
| Location/Qualifiers | | | |
| 1..2048 | | | |
| /organism="unidentified" | | | |
| /mol_type="genomic DNA" | | | |
| /db_xref="taxon:32644" | | | |
| FEATURES | | | |
| source | | | |
| ORIGIN | | | |
| Query Match | | | |
| Best Local Similarity | | | |
| Matches 2047; Conservative | | | |
| 0; Mismatches | | | |
| 1; Indels | | | |
| 4; Gaps | | | |
| 1; | | | |
| QY | 1 | CATGTGTGCCCTACAGCACATAGGGCTGTGGTTGAGAGAGCAGAGCTGCTCTGA | 60 |
| Db | 1 | CATGTGTGCCCTACAGCACATAGGGCTGTGGTTGAGAGAGCAGAGCTGCTCTGA | 60 |
| QY | 61 | CTTCTCTTCTTTTGACCTGTTGTATAAAGAGTAGAATAATTTTAAAAAGCTCGAA | 120 |
| Db | 61 | CTTCTCTTCTTTTGACCTGTTGTATAAAGAGTAGAATAATTTTAAAAAGCTCGAA | 120 |
| QY | 121 | TACTAACTCTCTCACAACTTCCGCTCTTTTCCAAACACCTTTTAACTTTTACT | 180 |
| Db | 121 | TACTAACTCTCTCACAACTTCCGCTCTTTTCCAAACACCTTTTAACTTTTACT | 180 |
| QY | 181 | TCTCATTTCTACTCCACTTCTTTTGGCTATAAGCAAGAAATTCACCTTCTTTTAAAGCTAACCCA | 240 |
| Db | 181 | TCTCATTTCTACTCCACTTCTTTTGGCTATAAGCAAGAAATTCACCTTCTTTTAAAGCTAACCCA | 240 |
| QY | 241 | AACGGCTCAATAAAGATCAATCAATAATGTATCTTTCAATTTTAGGATAAACAATACGT | 300 |

Db 241 AACGGCCTCAATAAAGATCAATTCATAAATGTATCTTCAATTTTAGGATAACAATACGT 300
Qy 301 GAACAGGGTATATTTTAAACGTGTCAACAAATCTAATAATTTTACCTGCCGGTGAACA 360
Db 301 GAACAGGGTATATTTTAAACGTGTCAACAAATCTAATAATTTTACCTGCCGGTGAACA 360
Qy 361 CCCTCTTCCAAGATAATATATTTTAAATTTTGTAGCTCCCTTTTAAACCAATTCGCAATGC 420
Db 361 CCCTCTTCCAAGATAATATATTTTAAATTTTGTAGCTCCCTTTTAAACCAATTCGCAATGC 420
Qy 421 AGGACGACTTAGGTGAATACACATTTGTAGTCTTTTAAACCAAGAAACAAGTGGTTC 480
Db 421 AGGACGACTTAGGTGAATACACATTTGTAGTCTTTTAAACCAAGAAACAAGTGGTTC 480
Qy 481 ATGCTCAGCCATCAAAATTTGAACAAACCCGACACACACTCTATCCACGTACTATCTTT 540
Db 481 ATGCTCAGCCATCAAAATTTGAACAAACCCGACACACACTCTATCCACGTACTATCTTT 540
Qy 541 TGGCCGAATGCTTCTCAAAATGTTTTTATATGTATATATGTTTACACTTTACAAGAGGATATTCGT 600
Db 541 TGGCCGAATGCTTCTCAAAATGTTTTTATATGTATATATGTTTACACTTTACAAGAGGATATTCGT 600
Qy 601 AAAATTCCTGTTTAAACAGTTTGTATATATATATGTTTACACTTTACAAGAGGATATTCGT 660
Db 601 AAAATTCCTGTTTAAACAGTTTGTATATATATATGTTTACACTTTACAAGAGGATATTCGT 660
Qy 661 AATACTTTTACAGACAGAGACTTAGGTCAAAATGGAGCTGTGTAACAGCCTAGACT 720
Db 661 AATACTTTTACAGACAGAGACTTAGGTCAAAATGGAGCTGTGTAACAGCCTAGACT 720
Qy 721 TGGTCACTGATAATAGATAATTTGTTAGTATATATATAGTAGGATCTCAATGATTAATA 780
Db 721 TGGTCACTGATAATAGATAATTTGTTAGTATATATATAGTAGGATCTCAATGATTAATA 780
Qy 781 ATTAGACTATTAATTAAGTTACTAATAAAGAGAGGTTAGTAACAGAAAGCAGGTA 840
Db 781 ATTAGACTATTAATTAAGTTACTAATAAAGAGAGGTTAGTAACAGAAAGCAGGTA 840
Qy 841 AAAACAGAGCTTGCTGCTGTGTGTTTGTAGTTGTTGTTG ---CATTTCTTTAAAAGTAATG 900
Db 841 AAAACAGAGCTTGCTGCTGTGTGTTTGTAGTTGTTGTTG ---CATTTCTTTAAAAGTAATG 896
Qy 901 TAACTGATCTAAAGCACAATAGAAATTTAGTACAGGTTTAAACCTTTTACAAGAAATTTATA 960
Db 901 TAACTGATCTAAAGCACAATAGAAATTTAGTACAGGTTTAAACCTTTTACAAGAAATTTATA 956
Qy 961 TTAACCAAGAAATCAATTTTAAACATGCTCTCGGCTGCTATTATAATAGGATCACTTAC 1020
Db 961 TTAACCAAGAAATCAATTTTAAACATGCTCTCGGCTGCTATTATAATAGGATCACTTAC 1016
Qy 1021 TGATCATCCATTTAAACCTTTGTTAAACAAATTCATAGAGATAAATATCTTACAATGAA 1080
Db 1021 TGATCATCCATTTAAACCTTTGTTAAACAAATTCATAGAGATAAATATCTTACAATGAA 1076
Qy 1081 AAGAAGCACAATGCTCTTTGAAAAAACAATAGGTACTCCCTCCGCTCTCTGAAATGT 1140
Db 1081 AAGAAGCACAATGCTCTTTGAAAAAACAATAGGTACTCCCTCCGCTCTCTGAAATGT 1136
Qy 1141 ATACATATGATTTGACACCGAGACTAAGAAATGATAAAGTATGATAGCTAATAAAG 1200
Db 1141 ATACATATGATTTGACACCGAGACTAAGAAATGATAAAGTATGATAGCTAATAAAG 1196
Qy 1201 AAAGAGAAAGAAAGTGGTAAAGTACGGGACCCCAATATATATATGATAGATTAG 1260
Db 1201 AAAGAGAAAGAAAGTGGTAAAGTACGGGACCCCAATATATATATGATAGATTAG 1256
Qy 1261 AAAAGTAGTTCAAAAGTAGTGGGTGGGATTTTTTATATATATAAATAATTTACTATTTTG 1320
Db 1261 AAAAGTAGTTCAAAAGTAGTGGGTGGGATTTTTTATATATAAATAATTTACTATTTTG 1316
Qy 1321 AGAAAGTTTGAATGATAGAAATTTAGTGGGACATCCATAAAGAAAGTATAGAAAT 1380
Db 1321 AGAAAGTTTGAATGATAGAAATTTAGTGGGACATCCATAAAGAAAGTATAGAAAT 1376

Qy 1381 TAAATGGGACAGAGGGAGTAATACCTTTATGATATATAAATTTTGTGTTATTTGATTTCA 1440
Db 1377 TAAATGGGACAGAGGGAGTAATACCTTTATGATATATAAATTTTGTGTTATTTGATTTCA 1436
Qy 1441 TAAGATTATAAACTCTATGTTTATTAATGATAATAAATTTTAAATAATATATATTAATTT 1500
Db 1437 TAAGATTATAAACTCTATGTTTATTAATGATAATAAATTTTAAATAATATATATTAATTT 1496
Qy 1501 CTGATTAGTCGATTACCGCCTTTTAAATTTTAACTACTGAGTAATATGAATAAATCAG 1560
Db 1497 CTGATTAGTCGATTACCGCCTTTTAAATTTTAACTACTGAGTAATATGAATAAATCAG 1556
Qy 1561 TTATCTGAAAGCAAAATAATATCTTTGTAACACAGCGTTCGGTCAAAATGGGAAGTTCAATG 1620
Db 1557 TTATCTGAAAGCAAAATAATATCTTTGTAACACAGCGTTCGGTCAAAATGGGAAGTTCAATG 1616
Qy 1621 TGTATTCAATAGTTTAAATAAAGTAAATTTTAAATTAATTTTGTGTTTCTGATTTTCTGAGA 1680
Db 1617 TGTATTCAATAGTTTAAATAAAGTAAATTTTAAATTAATTTTGTGTTTCTGAGA 1676
Qy 1681 AATTTAAATAAATAATTTAGAGCATGGAGTTCACGGGCATCATTTAGAGCAGCACTAGACT 1740
Db 1677 AATTTAAATAAATAATTTAGAGCATGGAGTTCACGGGCATCATTTAGAGCAGCACTAGACT 1736
Qy 1741 GTTTGAACAATGATGTCGGGTGATCATCTATGACCTTTCAACTCAAACTAGTGAATAAT 1800
Db 1737 GTTTGAACAATGATGTCGGGTGATCATCTATGACCTTTCAACTCAAACTAGTGAATAAT 1796
Qy 1801 GCATTTCTAGATAATCATCTTTCAAAATTTCAACAAACACAGCTTTAACTTTCTTTCAACG 1860
Db 1797 GCATTTCTAGATAATCATCTTTCAAAATTTCAACAAACACAGCTTTAACTTTCTTTCAACG 1856
Qy 1861 GATTGGAATCCTTTCTAAACTTTTAAATAAATAAATAAATGTCATTTTGTAAATTTTATC 1920
Db 1857 GATTGGAATCCTTTCTAAACTTTTAAATAAATAAATAAATGTCATTTTGTAAATTTTATC 1916
Qy 1921 AACACCTCAACATGATGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTACTATCAT 1980
Db 1917 AACACCTCAACATGATGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTACTATCAT 1976
Qy 1981 CACATCAATCTTACACCAACACCTTGAGCTTAATTTTCTACTTATTTCTCAGCAATAAC 2040
Db 1977 CACATCAATCTTACACCAACACCTTGAGCTTAATTTTCTACTTATTTCTCAGCAATCAC 2036
Qy 2041 ATTCTAAATATC 2052
Db 2037 ATTCTAAATATC 2048

RESULT 6
AR076817
LOCUS AR076817 2042 bp DNA linear PAT 30-AUG-2000
DEFINITION Sequence 2 from patent US 5959176.
ACCESSION AR076817
VERSION AR076817.1 GI:10003563
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2042)
AUTHORS Torikai,S. and Oeda,K.
TITLE Plant promoter and utilization thereof
JOURNAL Patent: US 5959176-A 2 28-SEP-1999;
FEATURES
 source 1..2042
 organism="unknown"
 mol_type="unassigned DNA"
ORIGIN

Query Match 11.8%; Score 241.2; DB 6; Length 2042;
Best Local Similarity 72.8%; Pred. No. 3.7e-28;
Matches 367; Conservative 0; Mismatches 128; Indels 9; Gaps 4;

1554 AATCAGTATCTGAAAGCAATAATATCTTTGTAACACAGCG----TTCCGGTCAAATG 1609
|||||
1543 AATATTATCTGAATGATACATCTTTGTAACAAACCTGGCCAAATAGACCATAA 1602
|||||
1610 GGAAGTTCATGTCATTAATAGTTTAAATATAAAGTAAATTTTAAATTAATGTTATT 1669
|||||
1603 CCAAGTTCACGTGATTTCTAAATGTTAAATACTAACATGAGTATTTCTTT--TTCAAGGT 1660
|||||
1670 TTTGTTTCAGAAATTTAAATAAATATTAGAGCATGGAAAGTTACCGGGCATCATTTGAGC 1729
|||||
1661 ATAAGTTAAATCTTCAATCAATTAATTTAAATTTGGACATTAATTGAGCAACTTTATGCC 1720
|||||
1730 AGCACTAGACTGTTTGAACCAATGATGTCGGGTGATCATCTATGACCTTTCAACTCAAAAC 1789
|||||
1721 CAGTGTATGTTTAAACAAAGTGTGTCGGGTGATATTTATGACCTTTCAACTCAAGC 1780
|||||
1790 TAGTGAAAT--AATGCAATCTAGAAATACATCTTTTCAAAATTTCAACAAACACAGCTTTAACT 1848
|||||
1781 TAGCCAGTGAATGCTTTCTAGAAATATATCTTTTGAATTTTCAACAAACACAGCACTAACT 1840
|||||
1849 TTTCTTTCAACGGATTTGGAATCTTTTCTAAATCTTTTAAATTTAAATAAATAATGCAATTT 1908
|||||
1841 TTTCTTTTAAACAGATTAGAAATCGTTTCTTAAACTTTTAAATTT--AAAAAATACATTTACT 1898
|||||
1909 GTAATATTTTCAACACTCAACATTTGATGTCGGGTGATCTATAAATAGGTGCTCTTGGTG 1968
|||||
1899 ATAATATTTTCAACACTCAACATTTGATGTCGGGTGATCTATAAATAGGTGCTCTTGGTG 1958
|||||
1969 CTCTACTATCATCATCAATCTTTACACCAAACTTTGAGCTTAATTTTCTACTTATT 2028
|||||
1959 CTCTACTATCATCATCAATCTTTCCAGCAAACTTTGAGCTTAATTTTCTACTTATT 2018
|||||
2029 CTGAGCAATCAATCTTCAATATC 2052
2019 TTTAGCAAAAACATCTTAAAGGTC 2042

RESULT 7
E15125
LOCUS E15125 2042 bp DNA linear PAT 28-JUL-1999
DEFINITION Promoter.
ACCESSION E15125
VERSION E15125.1 GI:5709808
KEYWORDS JP 1998052273-A/2.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 2042)
AUTHORS Torikai,T. and Oita,K.
TITLE VEGETABLE PROMOTER AND ITS USE
JOURNAL Patent: JP 1998052273-A 2 24-FEB-1998;
SUMITOMO CHEM CO LTD
COMMENT OS Daucus carota L. (carrot)
PN JP 1998052273-A/2
PD 24-FEB-1998
PF 12-AUG-1996 JP 1996212680
PI TORIKAI TOSHIMI, OITA KENJI
PC C12N15/09,A01H5/00,C07H21/04,C07K14/415,C12N1/21,C12N5/10; CC
strandedness: Single;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No; Location/Qualifiers
FH Key
FT source 1..2042 /organism='Daucus carota L.'
FT promoter 1<..<2042.
FT Location/Qualifiers
1..2042
/organism='unidentified'
/mol_type='genomic DNA'
/db_xref='taxon:32644'

ORIGIN

Query Match 11.8%; Score 241.2; DB 6; Length 2042;
Best Local Similarity 72.8%; Pred. No. 3.7e-28;
Matches 367; Conservative 0; Mismatches 128; Indels 9; Gaps 4;
1554 AATCAGTATCTGAAAGCAATAATATCTTTGTAACACAGCG----TTCCGGTCAAATG 1609
|||||
1543 AATATTATCTGAATGATACATCTTTGTAACAAACCTGGCCAAATAGACCATAA 1602
|||||
1610 GGAAGTTCATGTCATTAATAGTTTAAATATAAAGTAAATTTTAAATTAATGTTATT 1669
|||||
1603 CCAAGTTCACGTGATTTCTAAATGTTAAATACTAACATGAGTATTTCTTT--TTCAAGGT 1660
|||||
1670 TTTGTTTCAGAAATTTAAATAAATATTAGAGCATGGAAAGTTACCGGGCATCATTTGAGC 1729
|||||
1661 ATAAGTTAAATCTTCAATCAATTAATTTAAATTTGGACATTAATTGAGCAACTTTATGCC 1720
|||||
1730 AGCACTAGACTGTTTGAACCAATGATGTCGGGTGATCATCTATGACCTTTCAACTCAAAAC 1789
|||||
1721 CAGTGTATGTTTAAACAAAGTGTGTCGGGTGATATTTATGACCTTTCAACTCAAGC 1780
|||||
1790 TAGTGAAAT--AATGCAATCTAGAAATACATCTTTTCAAAATTTCAACAAACACAGCTTTAACT 1848
|||||
1781 TAGCCAGTGAATGCTTTCTAGAAATATATCTTTTGAATTTTCAACAAACACAGCACTAACT 1840
|||||
1849 TTTCTTTCAACGGATTTGGAATCTTTTCTAAATCTTTTAAATTTAAATAAATAATGCAATTT 1908
|||||
1841 TTTCTTTTAAACAGATTAGAAATCGTTTCTTAAACTTTTAAATTT--AAAAAATACATTTACT 1898
|||||
1909 GTAATATTTTCAACACTCAACATTTGATGTCGGGTGATCTATAAATAGGTGCTCTTGGTG 1968
|||||
1899 ATAATATTTTCAACACTCAACATTTGATGTCGGGTGATCTATAAATAGGTGCTCTTGGTG 1958
|||||
1969 CTCTACTATCATCATCAATCTTTACACCAAACTTTGAGCTTAATTTTCTACTTATT 2028
|||||
1959 CTCTACTATCATCATCAATCTTTCCAGCAAACTTTGAGCTTAATTTTCTACTTATT 2018
|||||
2029 CTGAGCAATCAATCTTCAATATC 2052
2019 TTTAGCAAAAACATCTTAAAGGTC 2042

RESULT 8

AR076816
LOCUS AR076816 247 bp DNA linear PAT 30-AUG-2000
DEFINITION Sequence 1 from patent US 5959176.
ACCESSION AR076816
VERSION AR076816.1 GI:10003562
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 247)
AUTHORS Torikai,S. and Oeda,K.
TITLE Plant promoter and utilization thereof
JOURNAL Patent: US 5959176-A 1 28-SEP-1999;
FEATURES
source 1..247
/organism='unknown'
/mol_type='unassigned DNA'

ORIGIN

Query Match 9.8%; Score 196.6; DB 6; Length 247;
Best Local Similarity 89.6%; Pred. No. 4.9e-21;
Matches 223; Conservative 0; Mismatches 24; Indels 2; Gaps 1;
1804 TTCTAGAAATACATCTTTTCAAAATTTCAACAAACACAGCTTTAACTTTTCTTTCAACGGAT 1863
1 TTCTAGAAATATATCTTTTGAATTTTCAACAAACACAGCACTAACTTTTCTTTTAAACAGAT 60
1864 TGGAAATCCTTTTCTAAACTTTTAAATATAAATAATGCAATTTATGTAATTTATCAAC 1923
|||||

```
Db 61 TAGAATCGTTCTTAACCTTTTAAATTT--AAAAATACATTACTATAATATTATCAAC 118
Qy 1924 ACCTCAACATGTAGTGTAGCGTACTATAAATAGGTGCTCTTGGTGTCTCTACTATCATCAC 1983
Db 119 ACCTCAACATTCATGTTAGCGTACTATAAATAGGTGCTCTTGGTGTCTCTACTATCATCAC 178
Qy 1984 ATCAATCTTACACACAAACCTTGAGCTTAATTTTCTTACTATTCTCAGCAATACATT 2043
Db 179 ATCAATCTTACACACAAACCTTGAGCTTAATTTTCTTACTAATTTTCTAGCAAAAACATT 238
Qy 2044 CTAATAATC 2052
Db 239 CTAAGGTC 247

RESULT 9
LOCUS E15124
DEFINITION Promoter.
ACCESSION E15124
VERSION E15124.1 GI:5709807
KEYWORDS JP 1998052273-A/1.
SOURCE unidentified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 247)
AUTHORS Torikai,T. and Oita,K.
TITLE VEGETABLE PROMOTER AND ITS USE
JOURNAL Patent: JP 1998052273-A 1 24-FEB-1998;
SUMITOMO CHEM CO LTD
COMMENT OS Daucus carota L. (carrot)
PN JP 1998052273-A/1
PD 24-FEB-1998
PF 12-AUG-1996 JP 1996212680
PI TORIKAI TOSHIMI, OITA KENJI
PC C12N15/09,A01H5/00,C07H21/04,C07K14/415,C12N1/21,C12N5/10; CC
scrandedness: Single;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
FH Key Location/Qualifiers
FH source 1..247
FT /organism="Daucus carota L."
FT /clones="pCR16GI-Xb"
FT promoter 1<..<247.
FT Location/Qualifiers
FEATURES
source 1..247
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
ORIGIN
Query Match 9.6%; Score 196.6; DB 6; Length 247;
Best Local Similarity 89.6%; Pred. No. 4.9e-21;
Matches 223; Conservative 0; Mismatches 24; Indels 2; Gaps 1;
Qy 1804 TTCTAGAATACATCTTTTCAAAATTTCAACAACACAGCTTTAACTTTTCTTCAACGGAT 1863
Db 1 TTCTAGAATATATCTTTTGAATTTTCAACAACACAGCACTAACTTTTCTTTTAACAGAT 60
Qy 1864 TGAATCCTTTCTTAACCTTTTAAAAATAAAAAATGCAATTTGTAAATATTATCAAC 1923
Db 61 TAGAATCGTTCTTAACCTTTTAAATTT--AAAAATACATTACTATAATATTATCAAC 118
Qy 1924 ACCTCAACATGTAGTGTAGCGTACTATAAATAGGTGCTCTTGGTGTCTCTACTATCATCAC 1983
Db 119 ACCTCAACATTCATGTTAGCGTACTATAAATAGGTGCTCTTGGTGTCTCTACTATCATCAC 178
Qy 1984 ATCAATCTTACACACAAACCTTGAGCTTAATTTTCTTACTATTCTCAGCAATACATT 2043
Db 179 ATCAATCTTACACACAAACCTTGAGCTTAATTTTCTTACTAATTTTCTAGCAAAAACATT 238
```

```
Qy 2044 CTAATAATC 2052
Db 239 CTAAGGTC 247

RESULT 10
LOCUS AR146852
DEFINITION Sequence 2 from patent US 6218598.
ACCESSION AR146852
VERSION AR146852.1 GI:15110041
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 246)
AUTHORS Ishige,F., Nishikawa,S. and Oeda,K.
TITLE Plant promoter
JOURNAL Patent: US 6218598-A 2 17-APR-2001;
FEATURES
source 1..246
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
Query Match 9.5%; Score 195.6; DB 6; Length 246;
Best Local Similarity 89.5%; Pred. No. 7e-21;
Matches 222; Conservative 0; Mismatches 24; Indels 2; Gaps 1;
Qy 1805 TCTAGAATACATCTTTTCAAAATTTCAACAACACAGCTTTAACTTTTCTTCAACGGATT 1864
Db 1 TCTAGAATATATCTTTTGAATTTTCAACAACACAGCACTAACTTTTCTTTTAACAGATT 60
Qy 1865 GGAATCCTTTCTAAACTTTTAAAAATAAAAAATGCAATTTGTAAATATTATCAACA 1924
Db 61 AGAATCGTTTCTTAACCTTTTAAATTT--AAAAATACATTACTATAATATTATCAACA 118
Qy 1925 CCTCAACATGTAGTGTAGCGTACTATAAATAGGTGCTCTTGGTGTCTCTACTATCATCAC 1984
Db 119 CCTCAACATTCATGTTAGCGTACTATAAATAGGTGCTCTTGGTGTCTCTACTATCATCAC 178
Qy 1985 TCAATCTTACACCACAAACCTTGAGCTTAATTTTCTTACTATTCTCAGCAATAACATTC 2044
Db 179 TCAATCTTCCAGCACAAACCTTGAGCTTAATTTTCTTACTAATTTTCTAGCAAAAACATTC 238
Qy 2045 TAAATATC 2052
Db 239 TAAAGGTC 246

RESULT 11
LOCUS E55065
DEFINITION Plant promoter.
ACCESSION E55065
VERSION E55065.1 GI:18625251
KEYWORDS JP 2000083679-A/2.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 246)
AUTHORS Ishige,I., Nishikawa,S. and Oeda,K.
TITLE Plant promoter
JOURNAL Patent: JP 2000083679-A 2 28-MAR-2000;
SUMITOMO CHEM CO LTD
COMMENT OS Daucus carota L.
PN JP 2000083679-A/2
PD 28-MAR-2000
PF 12-JUL-1999 JP 1999197240
PR IKUHARU ISHIGE, SATOMI NISHIKAWA, KENJI OEDA
PC C12N15/09,A01H5/00,C12N1/21,C12N5/10; ((C12N15/09,C12R1:91)), PC
((C12N5/10,C12R1:91)), C12N15/00,C12N5/00, C12N15/00,C12R1:91), PC
```

| | |
|--|---|
| (C12NS/00,C12R1:91) | |
| CC | Key |
| FT | source |
| FT | Location/Qualifiers |
| FT | 1..246 |
| FT | /organism="Daucus carota L." |
| FEATURES | |
| source | Location/Qualifiers |
| 1..246 | /organism="unidentified" |
| | /mol_type="genomic DNA" |
| | /db_xref="taxon:32644" |
| ORIGIN | |
| Query Match 9.5%; Score 195.6; DB 6; Length 246; | |
| Best Local Similarity 89.5%; Pred. No. 7e-21; | |
| Matches 22; Conservative 0; Mismatches 24; Indels 2; Gaps 1; | |
| Qy | 1805 TCTAGATACATCTTTTCAAAATTTCAACAAACACAGCTTTAACTTTCTTTTCAACGATT 1864 |
| Db | 1 TCTAGAAATATATCTTTTGAATTTTCAACAAACACAGCACTAACTTTCTTTTAAACGATT 60 |
| Qy | 1865 GGAATCCTTTTCTAACTTTTAAATAAAAAAATGCATTATTGTAATATTTTATCAACA 1924 |
| Db | 61 AGAATCGTTTCTTAACTTTTAAATTT--AAAAATACATTACTATTAATATTTATCAACA 118 |
| Qy | 1925 CCTCAACATTGATGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCTACTATCATCACAC 1984 |
| Db | 119 CCTCAACATTCAAGTTAGCGTACTATAAATAGGTGCTCTTGGTGCTCTACTATCATCACAC 178 |
| Qy | 1985 TCAATCTTACACCAACAACTTGAGCTTAATTTTCTTACTTATTTCTCAGCAATAACATTC 2044 |
| Db | 179 TCAATCTTCCAGCACAAACCTTGAGCTTAATCTTCTACTAATTTTGTAGCAAAACATTC 238 |
| Qy | 2045 TAAATATC 2052 |
| Db | 239 TAAAGGTC 246 |
| RESULT 12 | |
| DCA18706/c | |
| LOCUS | DCA18706 8005 bp DNA linear PLN 10-FEB-1999 |
| DEFINITION | Daucus carota Inv*Dc5 gene. |
| ACCESSION | Y18706 |
| VERSION | Y18706.1 GI:4454114 |
| KEYWORDS | beta-fructofuranosidase; Inv*Dc5 gene; invertase; isoform. |
| SOURCE | Daucus carota (carrot) |
| ORGANISM | Daucus carota |
| Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; campanulids; Apiales; Apiaceae; Apioidae; Scandiceae; Daucinae; Daucus. | |
| REFERENCE | 1 |
| AUTHORS | Sturm,A. |
| TITLE | Molecular characterisation and functional analysis of sucrose-cleaving enzymes in carrot (Daucus carota L.) |
| JOURNAL | J. Exp. Bot. 47, 1187-1192 (1996) |
| REFERENCE | 2 (bases 1 to 8005) |
| AUTHORS | Sturm,A. |
| TITLE | Direct Submission |
| JOURNAL | Submitted (27-JAN-1999) A. Sturm, Friedrich Miescher-Institut, Maulbeerstr. 66, CH-4058 Basel, SWITZERLAND |
| COMMENT | Related sequence X67163. |
| FEATURES | Location/Qualifiers |
| source | 1..8005 |
| | /organism="Daucus carota" |
| | /mol_type="genomic DNA" |
| | /cultivar="Nantaise" |
| | /db_xref="taxon:4039" |
| gene | 1..8005 |
| | /genes="Inv*Dc5" |
| 5'UTR | 1..2866 |
| CDS | join(2767..3160,3724..3732,4656..5515,5605..5766,5869..6104,6190..6277,6384..6587) |
| /gene="Inv*Dc5" | |
| /EC_number="3.2.1.26" | |
| /note="soluble acid invertase" | |
| /codon_start=1 | |
| /product="beta-fructofuranosidase, isoform II" | |
| /protein_id="CAA77266.1" | |
| /db_xref="GI:4454115" | |
| /db_xref="GOA:Q42722" | |
| /db_xref="UniProt/TREMBL:Q42722" | |
| /translation="MEHPITISHYTELPDGEHSPSLTTNTAEQSSRRSLTFLVLLFS SIIAACLVMTGTVLPFNSGNEAVEKSTVVPETEVEVAPRGVAEGVSMKSPRRPALNAE PPANFPWNSNVLISWQRSSFHQPNQNMNDPNGPLFYKGYHLFYQNPDPGATWGNKI VMGHAVSSDLIIHWKHLFVAMVTDHWYDVANGVTGSATILPDGQIVMLYTGSTNESQV LNKLTGSLAVYDVTDFKNTLLDQVLHAVHGTGMWECVDFVPVSKFGENGLDTSFGDVG VHKWAKASDDDRNDYYAIGTYDVPVSGKWVPDNPDLVDGIGLRYDGIYIYASKTFYDS NKRRRLVSWIKETDSEISDRKGMASVQGIPTRIILFDPKTKGSLNLLQWPVEEVNKKURL NKTVFNVINTGAVLPLEIGSGQLDITAEPEVDKESLERVOETNEVYDCKNKGSSS GRGALGPFGLLILADKDLSEQTPVIFYIAKSGGNLRTFFCADHSRSSKAVDVKIY GSVPVLRGEKLTMRILVDHSIVESFSGGRTCTISRVYPTKAIYNNAKVFLFNNAE ARIIASLNTWQMTAQROTHFADLVI" | |
| exon | 2767..3160 |
| | /gene="Inv*Dc5" |
| | /number=1 |
| intron | 3161..3723 |
| | /gene="Inv*Dc5" |
| | /number=1 |
| exon | 3724..3732 |
| | /gene="Inv*Dc5" |
| | /number=2 |
| intron | 3733..4655 |
| | /gene="Inv*Dc5" |
| | /number=2 |
| exon | 4656..5515 |
| | /gene="Inv*Dc5" |
| | /number=3 |
| intron | 5516..5604 |
| | /gene="Inv*Dc5" |
| | /number=3 |
| exon | 5605..5766 |
| | /gene="Inv*Dc5" |
| | /number=4 |
| intron | 5767..5868 |
| | /gene="Inv*Dc5" |
| | /number=4 |
| exon | 5869..6104 |
| | /gene="Inv*Dc5" |
| | /number=5 |
| intron | 6105..6189 |
| | /gene="Inv*Dc5" |
| | /number=5 |
| exon | 6190..6277 |
| | /gene="Inv*Dc5" |
| | /number=6 |
| intron | 6278..6383 |
| | /gene="Inv*Dc5" |
| | /number=6 |
| exon | 6384..6587 |
| | /gene="Inv*Dc5" |
| | /number=7 |
| 3'UTR | 6588..8005 |
| | /gene="Inv*Dc5" |
| ORIGIN | |
| Query Match 8.5%; Score 174; DB 8; Length 8005; | |
| Best Local Similarity 81.7%; Pred. No. 8.2e-18; | |
| Matches 201; Conservative 0; Mismatches 45; Indels 0; Gaps 0; | |
| Qy | 17 CACATAGGCGCTGTTTGGTTGAGAGAGCAGAGCTGCTCTCTCTCTCTCTTTGA 76 |
| Db | 1285 CACTATGGGGGTGTTTGTTCGCCGGAAGTAGGAGCTGCTCGCTCTCTCTCTGA 1226 |
| Qy | 77 CCTGTTTGTATAAAGAAGTAGAATATTTTAAAAAGCTGCGAATACTACTACTCTCTC 136 |

```
Db 1225 CCGCTTGTGTAATAAGCAGACATTTTAAAAAGTTGAAATGCTAGCTTCTCTC 1166
Qy 137 ACAATTCCTCGCTCTTTTCCAAACACTTTATTAACCTTTTATCTCTCTCAATTCCTCTCA 196
Db 1165 ACAGTTCCTGCTCTTTTCCAAACACTTTATTAATTTATTTACTTCTCACTTCTACTCTA 1106
Qy 197 CTTCTTTGCTATAGCAGAAATCACTCTTTTAAAGTAAACCCAAACGGCTCAATAAAA 256
Db 1105 CTTTCTAAATTAAGTAAGAAATCACTCTTTTAAAGCTAAACCCAAACGGCCCTATGACT 1046
Qy 257 GATCAT 262
Db 1045 GACCT 1040

RESULT 13
BD188672
LOCUS BD188672 2831 bp DNA linear PAT 17-JUL-2003
DEFINITION Promotor and use thereof.
ACCESSION BD188672
VERSION BD188672.1 GI:32998411
KEYWORDS JP 2003000252-A/1.
SOURCE unidentified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 2831)
AUTHORS Nishikawa,S. and Oeda,K.
TITLE Promotor and use thereof.
JOURNAL Patent: JP 2003000252-A 1 07-JAN-2003;
SUMITOMO CHEMICAL CO LTD
COMMENT OS Daucus carota L. (carrot)
PN JP 2003000252-A/1
PD 07-JAN-2003
PF 31-MAY-2001 JP 2001164069
PI SATOMI NISHIKAWA,KENJI OEDA
PC C12N15/09,A01H5/00,C12N1/19,C12N1/21,C12N5/10,C12N15/
PC 00,C12N5/00
CC Promotor and use thereof
FH Key Location/Qualifiers
FT promoter (1)..(2831).
FEATURES
source
1..2831
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
ORIGIN
Query Match 7.9%; Score 162.4; DB 6; Length 2831;
Best Local Similarity 82.5%; Pred. No. 6.4e-16;
Matches 196; Conservative 0; Mismatches 41; Indels 1; Gaps 1;
Qy 22 AGGCGCTGTTGGTTGAGAGAGCAGAGCTGCTTCTGACTTCTTCTCTTCTTCTGACCTGT 81
Db 1716 AGGCGCTGTTGGTTATGGAATCAGAGCTGCTTCTGACTTCTGCTCTCTCTCTTTT-ACCGGT 1774
Qy 82 TTGTATAAAGAGTAGAATAATTTTAAAAAGCTGCGAATACTTCTCTCTCTCAAC 141
Db 1775 TTGTATAAAGAGCAGACATTTTAAAGAGCTGAGAATGCTAGTTTCTCTCTCAGC 1834
Qy 142 TTGCGCTCTTTTCCAAACACTTTTAACTTTTACTTCTCTCTCTCTCTCTCTCT 201
Db 1835 TTCTGCTCTTTTCCAAACACTTTTCAACTTACCTTCTCTCTCTCTCTCTCTCT 1894
Qy 202 TTGCTATAAGCAAGAAATCACTTCTTTTAAAGCTAAACCCAAACGGCTCAATAAAGATCA 261
Db 1895 TTACTATAAGCAAGAGTCAATTTCTTTTAAATTAACCCAAACGGCCCTAAGTAATTGA 1954

RESULT 14
BD188674
LOCUS BD188674 2865 bp DNA linear PAT 17-JUL-2003
DEFINITION Promotor and use thereof.
```

```
ACCESSION BD188674
VERSION BD188674.1 GI:32998413
KEYWORDS JP 2003000252-A/3.
SOURCE unidentified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 2865)
AUTHORS Nishikawa,S. and Oeda,K.
TITLE Promotor and use thereof
JOURNAL Patent: JP 2003000252-A 3 07-JAN-2003;
SUMITOMO CHEMICAL CO LTD
COMMENT OS Daucus carota L. (carrot)
PN JP 2003000252-A/3
PD 07-JAN-2003
PF 31-MAY-2001 JP 2001164069
PI SATOMI NISHIKAWA,KENJI OEDA
PC C12N15/09,A01H5/00,C12N1/15,C12N1/19,C12N1/21,C12N5/10,C12N15/
PC 00,C12N5/00
CC Promotor and use thereof
FH Key Location/Qualifiers
FT promoter (1)..(2865).
FEATURES
source
1..2865
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
ORIGIN
Query Match 7.9%; Score 162.4; DB 6; Length 2865;
Best Local Similarity 82.5%; Pred. No. 6.4e-16;
Matches 196; Conservative 0; Mismatches 41; Indels 1; Gaps 1;
Qy 22 AGGCGCTGTTGGTTGAGAGAGCAGAGCTGCTTCTGACTTCTTCTTCTTCTTCTGACCTGT 81
Db 1716 AGGCGCTGTTGGTTATGGAATCAGAGCTGCTTCTGACTTCTGCTCTCTCTTTT-ACCGGT 1774
Qy 82 TTGTATAAAGAGTAGAATAATTTTAAAAAGCTGCGAATACTTCTCTCTCTCAAC 141
Db 1775 TTGTATAAAGAGCAGACATTTTAAAGAGCTGAGAATGCTAGTTTCTCTCTCAGC 1834
Qy 142 TTGCGCTCTTTTCCAAACACTTTTAACTTTTACTTCTCTCTCTCTCTCTCTCTCT 201
Db 1835 TTCTGCTCTTTTCCAAACACTTTTCAACTTACCTTCTCTCTCTCTCTCTCTCTCT 1894
Qy 202 TTGCTATAAGCAAGAAATCACTTCTTTTAAAGCTAAACCCAAACGGCTCAATAAAGATCA 261
Db 1895 TTACTATAAGCAAGAGTCAATTTCTTTTAAATTAACCCAAACGGCCCTAAGTAATTGA 1954

RESULT 15
DARGCHS2
LOCUS DARGCHS2 4886 bp DNA linear PLN 14-APR-2000
DEFINITION Carrot gCHS2 gene for chalcone synthase.
ACCESSION D16255
VERSION D16255.1 GI:441168
KEYWORDS chalcone synthase.
SOURCE Daucus carota (carrot)
ORGANISM Daucus carota
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; campanulids; Apiales; Apiaceae; Apioidae; Scandiceae;
Daucinae; Daucus.
REFERENCE 1 (bases 1 to 4886)
AUTHORS Ozeki,Y., Davies,E. and Takeda,J.
TITLE Structure and expression of chalcone synthase gene in carrot
JOURNAL Plant Cell Physiol. 34, 1029-1037 (1993)
REFERENCE 2 (bases 1 to 4886)
AUTHORS Ozeki,Y.
TITLE Direct Submission
JOURNAL Submitted (14-MAY-1993) Yoshihiro Ozeki, College of Arts and
Sciences, The University of Tokyo, Department of Biology; Komaba,
Meguro-ku, Tokyo 153, Japan (Tel:03-3467-1171(ex.253)).
```

Search completed: March 14, 2005, 18:28:48
Job time : 9015 secs

COMMENT
Submitted (14-MAY-1993) to DDBJ by:
Yoshihiro Ozeki
Department of Biol. College of
Arts & Science, Univ. of Tokyo
3-8-1 Komaba, Meguro-ku
Tokyo 153
Japan
Phone: 03-3467-1171 x253
Fax: 03-3485-2904.

FEATURES
source
Location/Qualifiers
1. .4886

/organism="Daucus carota"
/mol_type="genomic DNA"
/cultivar="Kurodagozun"
/db_xref="taxon:4039"
/dev_stage="cultured cells"
1957. .1962

misc_feature
/note="TACPyAT motif"
2060. .2065

misc_feature
/note="G-box"
2286. .2291

misc_feature
/note="G-box"
2303. .2308

misc_feature
/note="TACPyAT motif"
2327. .2331

TATA_signal
2358. .4519

prim_transcript
/note="CHS2 mRNA and intron"
2358. .2607

exon
/number=1

gene
join(2430. .2607,3395. .4386)

CDS
/genes="gCHS2"
join(2430. .2607,3395. .4386)

/gene="gCHS2"
/EC_number="2.3.1.74"

/codon_start=1
/product="chalcone synthase"

/protein_id="BAA03784.1"
/db_xref="GI:565639"

/translation="MTVNEFRKAQRAEGPATVLAIGTATPPNCVDQSAVADYYFRIT
NSDKEPELKEKRMCEKMINTRYMHLTEDLLKQNPSCFENYASSLDARQDIVNEV
PLKGEAAARAIKEWQGPQSKITHLIFCTTSGVMPGADFRLTKLGLRPSVKRFNMY
QOCFAGTGVLRKADLAENKNARVLVVCSEITVITFRGPNDLHDSLVGOALFGDG
AGAVIVGSDPVIIEKPLPEIVSAQOTILPDSGADJGHLREVGTLFHLKQVPGILIS
KNIRKSLVFAFKPLGISDNNISIFWIAHFGGPAILDQVETELSLKPEKLKSTRQVLRDY
GNSSACVLFLIDEMRKASAKOGRITTBEGLDWGVLFGRPGGLTVETVVLHVSPT"

2608. .3394
/gene="gCHS2"

3395. .4519
/number=2

polyA_signal
4501. .4506

ORIGIN

Query Match 7.9%; Score 162.4; DB 8; Length 4886;
Best Local Similarity 82.5%; Pred. No. 5.8e-16;
Matches 198; Conservative 0; Mismatches 41; Indels 1; Gaps 1;

QY 11 CTACAGCACATAGGGCTGTTGGTTGAGAGAGCAGAGCTGCTTCTGACTTCTTCTTC 70

Db 1355 CTAGTAGATGTAGGGCTGTTGGATGTTGGAAGCAGAGCTGCTTCCGGCTTCTGCTTT 1414

QY 71 TTTTGACCTGTTGTATAAGAGATGAGAAATATTTTAAAAAGCTCGAATACTAACTTC 130

Db 1415 TCTTGACCCGTTTGGGTAAAGAGTAGAGCCTTTTAAAGAGTTGAGATAGCTTC 1474

QY 131 TCCTCACAACTTCG-CTTCTTTTCCAAACACATTTATTAACTTTTACTTCTCAATTC 189

Db 1475 TCCTCAGAGCTTCTGTTTTTTTCCAAACACATTTATTCAATTTACTTTTCACTTC 1534

QY 190 TACTCCACTCTTGTGTATAGCAGAGAAATCACTTCTTTTAAAGCTAACCCAAACGGCTC 249

Db 1535 TACTCCACTCTTGTAAATTTAAGCAAGAGTCACTTCTTTTAAACTAACCCAAACGGCTC 1594

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 14, 2005, 14:16:29 ; Search time 1121 Seconds
(without alignments)
10836.137 Million cell updates/sec

Title: US-09-806-197-1
Perfect score: 2052
Sequence: 1 cagtgtgcctacagcaca.....gcaataacattctaaatatc 2052

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004as:*
13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|---------------|---------------------|
| 1 | 2052 | 100.0 | 2052 | 3 AAA37959 | Aaa37959 Carrot pr |
| 2 | 2048.8 | 99.8 | 2052 | 3 AAA37961 | Aaa37961 Carrot pr |
| 3 | 2034.8 | 99.2 | 2056 | 3 AAA37964 | Aaa37964 Plasmid # |
| 4 | 2030.8 | 99.0 | 2048 | 3 AAA37962 | Aaa37962 Plasmid # |
| 5 | 2030.8 | 99.0 | 2048 | 3 AAA37963 | Aaa37963 Plasmid # |
| 6 | 239.6 | 11.7 | 2042 | 2 AAV15144 | AAV15144 New promo |
| 7 | 196.6 | 9.6 | 247 | 2 AAV15143 | AAV15143 New promo |
| 8 | 195.6 | 9.5 | 246 | 3 AAZ49611 | Aaz49611 Carrot CR |
| 9 | 194.4 | 9.5 | 196 | 12 ADP07499 | Adp07499 Carrot DN |
| 10 | 162.4 | 7.9 | 2831 | 10 ADC56759 | Adc56759 Carrot DN |
| 11 | 162.4 | 7.9 | 2865 | 10 ADC56761 | Adc56761 Carrot DN |
| c 12 | 119 | 5.8 | 140 | 3 AAZ49616 | Aaz49616 Oligonucle |
| 13 | 119 | 5.8 | 140 | 3 AAZ49615 | Aaz49615 Oligonucle |
| 14 | 99.4 | 4.8 | 8056 | 8 ABZ10246 | Abz10246 Haematopo |
| 15 | 91.8 | 4.5 | 8056 | 8 ABZ10100 | Abz10100 Haematopo |
| c 16 | 91.6 | 4.5 | 8056 | 8 ABZ10246 | Abz10246 Haematopo |
| c 17 | 88.6 | 4.3 | 851 | 3 AAA37960 | Aaa37960 Carrot te |
| c 18 | 88.6 | 4.3 | 8056 | 8 ABZ10100 | Abz10100 Haematopo |
| c 19 | 87.2 | 4.2 | 110000 | 13 ABD32968_6 | Continuation (7 of |
| 20 | 83 | 4.0 | 158001 | 12 ADL17884 | Adl17884 Human pho |

| | | | | | | |
|------|------|-----|--------|----|------------|---------------------|
| c 21 | 75 | 3.7 | 11745 | 6 | ABK28332 | Abk28332 DNA trans |
| c 22 | 71.2 | 3.5 | 8222 | 8 | ACF62816 | Acf62816 Colon can |
| c 23 | 71.2 | 3.5 | 11222 | 10 | ADB54318 | Adb54318 Pretreat |
| c 24 | 71 | 3.5 | 5930 | 6 | ABL32517 | Ab132517 Human imm |
| c 25 | 70.8 | 3.5 | 5286 | 13 | ADS89278 | Ads89278 Oligonucle |
| c 26 | 70.8 | 3.5 | 5286 | 13 | ADS89552 | Ads89552 Oligonucle |
| c 27 | 69.8 | 3.4 | 6352 | 6 | ABK31340 | Abk31340 Signal tr |
| c 28 | 69.8 | 3.4 | 6352 | 6 | ABL70563 | Ab170563 Chemicall |
| c 29 | 69.8 | 3.4 | 6352 | 6 | AAS61235 | Aas61235 Human gen |
| c 30 | 69.6 | 3.4 | 8222 | 8 | ACF62794 | Acf62794 Colon can |
| c 31 | 69.6 | 3.4 | 11222 | 10 | ADB54190 | Adb54190 Pretreat |
| c 32 | 69.4 | 3.4 | 16258 | 6 | ABK40038 | Abk40038 Human che |
| c 33 | 69.4 | 3.4 | 16258 | 6 | ABL70376 | Ab170376 Chemicall |
| c 34 | 69.2 | 3.4 | 6045 | 6 | ABK31541 | Abk31541 Signal tr |
| c 35 | 69.2 | 3.4 | 6045 | 6 | ABL70624 | Ab170624 Chemicall |
| c 36 | 68.2 | 3.3 | 15548 | 6 | ABL34155 | Ab134155 Human imm |
| c 37 | 68.2 | 3.3 | 110000 | 13 | ABD32968_6 | Continuation (7 of |
| c 38 | 68 | 3.3 | 11422 | 6 | ABK39936 | Abk39936 Human che |
| c 39 | 68 | 3.3 | 11422 | 6 | ABL32218 | Ab132218 Human imm |
| c 40 | 67.8 | 3.3 | 15732 | 4 | AAS45388 | Aas45388 Chemicall |
| c 41 | 67.8 | 3.3 | 15732 | 6 | ABK28233 | Abk28233 DNA trans |
| c 42 | 67.8 | 3.3 | 99764 | 13 | ABD32992 | Abd32992 Human can |
| c 43 | 67.6 | 3.3 | 5984 | 6 | ABQ66994 | Abq66994 Human ang |
| c 44 | 67.6 | 3.3 | 6050 | 6 | ABL34011 | Ab134011 Human imm |
| c 45 | 67.4 | 3.3 | 5286 | 13 | ADS89551 | Ads89551 Oligonucle |

ALIGNMENTS

RESULT 1
AAA37959
ID AAA37959 standard; DNA; 2052 BP.
XX
AC AAA37959;
XX
DT 18-AUG-2000 (first entry)
XX
DE Carrot promoter sequence #1.
XX
KW Carrot; promoter; terminator; transgenic plant; breeding; fertility; ds.
XX
OS Daucus carota.
XX
WO2000020613-A1.
PN
XX
PD 13-APR-2000.
XX
PF 28-SEP-1999; 99WO-JP005303.
XX
PR 02-OCT-1998; 98JP-00281124.
XX
PA (SUMO) SUMITOMO CHEM CO LTD.
XX
PI Nishikawa S, Oeda K;
XX
DR WPI; 2000-303791/26.
XX
PT New Plant promoters and terminators from Daucus carota L., useful in
PT plant breeding, for e.g. controlling fertilities of plants.
XX
PS Claim 1; Page 69-70; 81pp; English.
XX
CC This sequence represents a carrot promoter. The invention relates to
CC plant promoters and terminators from Daucus carota L. which are capable
CC of expressing a gene of interest in plants. The invention also includes a
CC chimeric gene characterized in that it comprises the promoter and a
CC desired gene linked to each other in the form capable of functioning. A
CC method of producing a transformant comprises introducing the promoter,
CC the chimeric gene or a vector comprising the promoter and a desired gene
CC or terminator sequence into a host cell. The plant promoters and
CC terminators are useful in plant breeding, for e.g. fertilities of plants
CC may be controlled by expressing, in the host cells, a sense or antisense

| | | | |
|----|---|------------------------|--|
| Dd | | 2041 ATTCTAAATATC 2052 | |
| | RESULT 2 | | |
| | AAA37961 | | |
| ID | AAA37961 standard; DNA; 2052 BP. | | |
| XX | | | |
| AC | AAA37961; | | |
| XX | | | |
| DT | 18-AUG-2000 (first entry) | | |
| DE | Carrot promoter sequence #2. | | |
| XX | | | |
| KW | Carrot; promoter; terminator; transgenic plant; breeding; fertility; ds. | | |
| OS | Daucus carota. | | |
| PN | WO200020613-A1. | | |
| PD | 13-APR-2000. | | |
| XX | | | |
| Pf | 28-SEP-1999; 99WO-JP005303. | | |
| PR | 02-OCT-1998; 98JP-00281124. | | |
| PA | (SUMO) SUMITOMO CHEM CO LTD. | | |
| PI | Nishikawa S, Oeda K; | | |
| XX | | | |
| DR | WFI; 2000-303791/26. | | |
| PT | New Plant promoters and terminators from Daucus carota L., useful in plant breeding, for e.g. controlling fertilities of plants. | | |
| PS | Claim 1; Page 78-79; 81pp; English. | | |
| CC | This sequence represents a carrot promoter. The invention relates to plant promoters and terminators from Daucus carota L. which are capable of expressing a gene of interest in plants. The invention also includes a chimeric gene characterized in that it comprises the promoter and a desired gene linked to each other in the form capable of functioning. A method of producing a transformant comprises introducing the promoter, the chimeric gene or a vector comprising the promoter and a desired gene or terminator sequence into a host cell. The plant promoters and terminators are useful in plant breeding, for e.g. fertilities of plants may be controlled by expressing, in the host cells, a sense or antisense gene of a male sterility related gene such as S-locus-specific RNase gene | | |
| SQ | Sequence 2052 BP; 736 A; 318 C; 317 G; 681 T; 0 U; 0 Other; | | |
| | Query Match 99.8%; Score 2048.8; DB 3; Length 2052; Best Local Similarity 99.9%; Pred. No. 0; Matches 2050; Conservative 0; Mismatches 2; Indels 0; Gaps 0; | | |
| Qy | 1 CATGTGTCCTACAGCACATAGGGCGTCTTTGGTTGAGAGAAGCAGAGCTGCTTCTGA 60 | | |
| Dd | 1 CATGTGTCCTACAGCACATAGGGCGTCTTTGGTTGAGAGAAGCAGAGCTGCTTCTGA 60 | | |
| Qy | 61 CTTCCTCTCTTTGACCTGTTTGTAATAAAGTAGATAATTTTTAAAAAGCTGCAGAA 120 | | |
| Dd | 61 CTTCCTCTCTTTGACCTGTTTGTAATAAAGTAGATAATTTTTAAAAAGCTGCAGAA 120 | | |
| Qy | 121 TACTTAACCTCTCTCACAACCTTCGCTCTTTTCCACAACCTTTATTAACTTTTTTACT 180 | | |
| Dd | 121 TACTTAACCTCTCTCACAACCTTCGCTCTTTTCCACAACCTTTATTAACTTTTTTACT 180 | | |
| Qy | 181 TCTCATTTCTACTCCACTCTTTTGTCTATAAGCAAGAAATCACTCTTTTAAGCTAACCCA 240 | | |
| Dd | 181 TCTCATTTCTACTCCACTCTTTTGTCTATAAGCAAGAAATCACTCTTTTAAGCTAACCCA 240 | | |
| Qy | 241 AACGGCCTCAATAAAGATCATTCATPAATATGATCTTTCAAATTTTAGATTAACATAGT 300 | | |
| Dd | 241 AACGGCCTCAATAAAGATCATTCATPAATATGATCTTTCAAATTTTAGATTAACATAGT 300 | | |


```
Db 661 |AATACCTTTTACGACACAGAGACTTTAGTCAAAATGGACCTGTGTAACAGCCTAGACT 720
Qy 721 TGGTCACCTGATAAATAGATAATTTGTAGTAAATAGTAGGATCTCAATAGCAATATAA 780
Db 721 TGGTCACCTGATAAATAGATAATTTGTAGTAAATAGTAGGATCTCAATAGCAATATAA 780
Qy 781 ATTAGAGCTATTAATTAAGTACTAAATAAAGAGAGGTTAGTAACAGAAAGCAGGTA 840
Db 781 ATTAGAGCTATTAATTAAGTACTAAATAAAGAGAGGTTAGTAACAGAAAGCAGGTA 840
Qy 841 AAAACAAGAGCTCTGCTGCTGTTAGTTAGTTGTTGAGCTCATTTCTTTAAAGTAATG 900
Db 841 AAAACAAGAGCTCTGCTGCTGTTAGTTAGTTGTTGAGCTCATTTCTTTAAAGTAATG 900
Qy 901 TAAACTGATCTAAAGCACAATAGAAATTTAGTACAGGTTTAAACCTTTTACAGAAATTTATA 960
Db 901 TAAACTGATCTAAAGCACAATAGAAATTTAGTACAGGTTTAAACCTTTTACAGAAATTTATA 960
Qy 961 TTAACGAAATATCAATTTTAAACATGCTCTCGCTGTCATATAAAGGATCACTTAC 1020
Db 961 TTAACGAAATATCAATTTTAAACATGCTCTCGCTGTCATATAAAGGATCACTTAC 1020
Qy 1021 TGATCATCCATTAACCTTTTAAACAAATTCATGAGTAAATATCTTACAATGAA 1080
Db 1021 TGATCATCCATTAACCTTTTAAACAAATTCATGAGTAAATATCTTACAATGAA 1080
Qy 1081 AAGAAGACATGCTCTTTGAAAACAAATAGTACTCCCTCCGCTCTCAATGT 1140
Db 1081 AAGAAGACATGCTCTTTGAAAACAAATAGTACTCCCTCCGCTCTCAATGT 1140
Qy 1141 ATAATATGAGTATGAGACAGGAGACTAAGAAAATGATATAAGTAATGTAGAGTAAAG 1200
Db 1141 ATACATATGAGTATGAGACAGGAGACTAAGAAAATGATATAAGTAATGTAGAGTAAAG 1200
Qy 1201 AAAGAGAAAGAAAGTGGTAAAGTACGGGACCCCAATATATATGATAGATTAG 1260
Db 1201 AAAGAGAAAGAAAGTGGTAAAGTACGGGACCCCAATATATATGATAGATTAG 1260
Qy 1261 AAAAGTAGTTGAAAGTAGTGGGTGGGATTTTATATATATAAATTTTACTATTTTG 1320
Db 1261 AAAAGTAGTTGAAAGTAGTGGGTGGGATTTTATATATATAAATTTTACTATTTTG 1320
Qy 1321 AGAAAGTTTGAATATAGAAATGAGTGGGACATCCATAAAGGAAAGTATAGAAT 1380
Db 1321 AGAAAGTTTGAATATAGAAATGAGTGGGACATCCATAAAGGAAAGTATAGAAT 1380
Qy 1381 TAAATGGACAGAGGAGTAATACCTTTATGATATATAAATTTTGTATTTGATTCA 1440
Db 1381 TAAATGGACAGAGGAGTAATACCTTTATGATATATAAATTTTGTATTTGATTCA 1440
Qy 1441 TAAGATTATAATCTATGTTATATATATATATATATATATATATATATATATAT 1500
Db 1441 TAAGATTATAATCTATGTTATATATATATATATATATATATATATATATATAT 1500
Qy 1501 CTGATTAGTCGATTACCGCTTTTATATATTTTACAATCTGAGTAATGATAAATCAG 1560
Db 1501 CTGATTAGTCGATTACCGCTTTTATATATTTTACAATCTGAGTAATGATAAATCAG 1560
Qy 1561 TTATCTGAAAGCAAAATATCTTTGTAAACAGCGTTCGGTCAATGGGAAGTTCATG 1620
Db 1561 TTATCTGAAAGCAAAATATCTTTGTAAACAGCGTTCGGTCAATGGGAAGTTCATG 1620
Qy 1621 TGTATTCATATAGTTTATATATAAAGTAAATTTTAAATTAATCTTATTTTGTTCAGA 1680
Db 1621 TGTATTCATATAGTTTATATATAAAGTAAATTTTAAATTAATCTTATTTTGTTCAGA 1680
Qy 1681 AATTAAATATAATTTATGAGCATGGGAAGTTACGGGCAATGAGCAGCACTAGACT 1740
Db 1681 AATTAAATATAATTTATGAGCATGGGAAGTTACGGGCAATGAGCAGCACTAGACT 1740
Qy 1741 GTTTGAACAATGTATGCGGTGTACATCTATGACCTTTCAACTCAACTAGTGAATAAT 1800
```

```
Db 1741 GTTTGAACAATGTATGCGGTGTACATCTATGACCTTTCAACTCAACTAGTGAATAAT 1800
Qy 1801 GCATT-----CTAGAATACATCTTTTCAAATTTCAACAAACACAGCTTTTAACTTTTCTTTC 1856
Db 1801 GCATTCTAGCTAGAAATACATCTTTTCAAATTTCAACAAACACAGCTTTTAACTTTTCTTTC 1860
Qy 1857 AACGGATTGGAATCCCTTTTCTAAACCTTTTAAATATAAATAAATAAATGCAATATTGTAATATT 1916
Db 1861 AACGGATTGGAATCCCTTTTCTAAACCTTTTAAATATAAATAAATAAATGCAATATTGTAATATT 1920
Qy 1917 TATCAACACCTCAACATTTAGTGTAGCTACTATAAATAGTGTCTCTTTGGTGTCTCTACTA 1976
Db 1921 TATCAACACCTCAACATTTAGTGTAGCTACTATAAATAGTGTCTCTTTGGTGTCTCTACTA 1980
Qy 1977 TCATCACATCAATCTTACACCAACACCTTTGAGCTTAATTTTCTACTTATTTCTACGAA 2036
Db 1981 TCATCACATCAATCTTACACCAACACCTTTGAGCTTAATTTTCTACTTATTTCTACGAA 2040
Qy 2037 TAACATTTCTAAATATC 2052
Db 2041 TCACATTTCTAAAGATC 2056

RESULT 4
AAA37962
ID AAA37962 standard; DNA; 2048 BP.
XX
AC AAA37962;
XX
DT 18-AUG-2000 (first entry)
XX
DE Plasmid #1 DNA sequence used in mutation of promoter sequence.
XX
KW Carrot; promoter; terminator; transgenic plant; breeding; fertility; ds.
XX
OS Daucus carota.
XX
PN W0200020613-A1.
XX
PD 13-APR-2000.
XX
PF 28-SEP-1999; 99WO-JP005303.
XX
PR 02-OCT-1998; 98JP-00281124.
XX
PA (SUMO ) SUMITOMO CHEM CO LTD.
XX
PI Nishikawa S, Oeda K;
XX
DR WPI; 2000-303791/26.
XX
PT New Plant promoters and terminators from Daucus carota L., useful in
PT plant breeding, for e.g. controlling fertilities of plants.
XX
PS Example 8; Page 71-73; 81pp; English.
XX
CC This sequence represents a plasmid sequence used in a method for
CC introducing a mutation into a carrot promoter. The invention relates to
CC plant promoters and terminators from Daucus carota L. which are capable
CC of expressing a gene of interest in plants. The invention also includes a
CC chimeric gene characterized in that it comprises the promoter and a
CC desired gene linked to each other in the form capable of functioning. A
CC method of producing a transformant comprises introducing the promoter,
CC the chimeric gene or a vector comprising the promoter and a desired gene
CC or terminators sequence into a host cell. The plant promoters and
CC terminators are useful in plant breeding, for e.g. fertilities of plants
CC may be controlled by expressing, in the host cells, a sense or antisense
CC gene of a male sterility related gene such as S-locus-specific RNase gene
XX
SQ Sequence 2048 BP; 735 A; 317 C; 316 G; 680 T; 0 U; 0 Other;

Query Match 99.0%; Score 2030.8; DB 3; Length 2048;
Best Local Similarity 99.7%; Pred. No. 0;
```


| | | | | | |
|----|---|----|------|--|------|
| XX | AAA37963; | QY | 361 | CGGTCTTCCAGTATATATATTTTAAATTTTGTAGCCTCCCTTTTAAACCAAAATTCGCATGC | 420 |
| AC | | DB | 361 | | |
| XX | 18-AUG-2000 (first entry) | | | CGGTCTTCCAGTATATATATTTTAAATTTTGTAGCCTCCCTTTTAAACCAAAATTC---GC | 416 |
| DT | | | | | |
| XX | Plasmid #2 DNA sequence used in mutation of promoter sequence. | QY | 421 | AGGACGACTTAGGTGAATACACATCTGTAAGTCTTTTAAACAAAGAACAGTGGTTC | 480 |
| DE | | | | | |
| XX | Carrot; promoter; terminator; transgenic plant; breeding; fertility; ds. | DB | 417 | AGGACGACTTAGGTGAATACACATCTGTAAGTCTTTTAAACAAAGAACAGTGGTTC | 476 |
| KW | | | | | |
| XX | Daucus carota. | QY | 481 | ATGCTCAGCCATCAAAATTTGACAAAACCCGACACAAACACTCTATCCAGTACTATATCTTT | 540 |
| OS | | DB | 477 | ATGCTCAGCCATCAAAATTTGACAAAACCCGACACAAACACTCTATCCAGTACTATATCTTT | 536 |
| XX | WO200020613-A1. | | | | |
| PN | 13-APR-2000. | QY | 541 | TGGCCGAATGCTCTCAGAAATGTTTTTATATGTAATAAATAAGTCCCATCCAAAGATAAGT | 600 |
| XX | | | | | |
| PD | 28-SEP-1999; 99WO-JP005303. | DB | 537 | TGGCCGAATGCTCTCAGAAATGTTTTTATATGTAATAAATAAGTCCCATCCAAAGATAAGT | 596 |
| XX | | | | | |
| PF | 02-OCT-1998; 98JP-00281124. | QY | 601 | AAAAATCCCGTTTAAACGAGTTTGTAAATATATATGTTTACCTTACAAAGAGGATATTCGT | 660 |
| XX | | | | | |
| PR | (SUMO) SUMITOMO CHEM CO LTD. | DB | 597 | AAAAATCCCGTTTAAACGAGTTTGTAAATATATATGTTTACCTTACAAAGAGGATATTCGT | 656 |
| XX | Nishikawa S, Oeda K; | | | | |
| PI | WPI; 2000-303791/26. | QY | 661 | AATACTTTTAGACGACAGAGACTTAGGTCAAAAATGGACGCTGGTGAACAGCCTAGACT | 720 |
| XX | | DB | 657 | AATACTTTTAGACGACAGAGACTTAGGTCAAAAATGGACGCTGGTGAACAGCCTAGACT | 716 |
| XX | New Plant promoters and terminators from Daucus carota L., useful in | QY | 721 | TGCTCACTGATTAATAGATTAATTTAGTATATATATAGTAGGATCTACATGACATATAA | 780 |
| PT | plant breeding, for e.g. controlling fertilities of plants. | DB | 717 | TGCTCACTGATTAATAGATTAATTTAGTATATATATAGTAGGATCTACATGACATATAA | 776 |
| XX | | | | | |
| PS | Example 8; Page 73-74; 81pp; English. | QY | 781 | ATTAGAGCTATTAAATTAAGTTACTATAAATAAGAGAGGTTAGTAAACAGAAAAGCAGGTA | 840 |
| XX | | DB | 777 | ATTAGAGCTATTAAATTAAGTTACTATAAATAAGAGAGGTTAGTAAACAGAAAAGCAGGTA | 836 |
| CC | This sequence represents a plasmid sequence used in a method for | QY | 841 | AAAAAAGAGCTTGTCTGT | 900 |
| CC | introducing a mutation into a carrot promoter. The invention relates to | | | | |
| CC | plant promoters and terminators from Daucus carota L. which are capable | DB | 837 | AAAAAAGAGCTTGTCTGT | 896 |
| CC | of expressing a gene of interest in plants. The invention also includes a | | | | |
| CC | chimeric gene characterized in that it comprises the promoter and a | QY | 901 | TAAATGATCTAAAGCACATAGAAATTTAGTACAGGTTAAACCTTTTACAAAGATTTATA | 960 |
| CC | desired gene linked to each other in the form capable of functioning. A | DB | 897 | TAAATGATCTAAAGCACATAGAAATTTAGTACAGGTTAAACCTTTTACAAAGATTTATA | 956 |
| CC | method of producing a transformant comprises introducing the promoter, | | | | |
| CC | the chimeric gene or a vector comprising the promoter and a desired gene | QY | 961 | TTAAACGAAAATCATTTTATAACATGCTCTCGGCTGTCTATTATAATAGGATCAGCTTAC | 1020 |
| CC | or terminator sequence into a host cell. The plant promoters and | DB | 957 | TTAAACGAAAATCATTTTATAACATGCTCTCGGCTGTCTATTATAATAGGATCAGCTTAC | 1016 |
| CC | terminators are useful in plant breeding, for e.g. fertilities of plants | | | | |
| CC | may be controlled by expressing, in the host cells, a sense or antisense | QY | 1021 | TGATCATCCATTAAAAACCTTGT | 1080 |
| CC | gene of a male sterility related gene such as S-locus-specific RNase gene | DB | 1017 | TGATCATCCATTAAAAACCTTGT | 1076 |
| XX | Sequence 2048 BP; 735 A; 317 C; 316 G; 680 T; 0 U; 0 Other; | | | | |
| SQ | | QY | 1081 | AGAAGGACAAATGTCTCTTTGAAAAAACAATAAGTAGTACTCCCTCCCTCTGAAATGT | 1140 |
| | | DB | 1077 | AGAAGGACAAATGTCTCTTTGAAAAAACAATAAGTAGTACTCCCTCCCTCTGAAATGT | 1136 |
| | | | | | |
| | | QY | 1141 | ATACATATGATTTGGACACGAGACTTAAAGAAAATGTATATAAGTAAATGTAGAGTAAAAAG | 1200 |
| | | DB | 1137 | ATACATATGATTTGGACACGAGACTTAAAGAAAATGTATATAAGTAAATGTAGAGTAAAAAG | 1196 |
| | | QY | 1201 | AAAAGAGAAAAGTGGGTAAAGTAGCGGGACCCCAATATATATAATTTGATAGATTTAG | 1260 |
| | | DB | 1197 | AAAAGAGAAAAGTGGGTAAAGTAGCGGGACCCCAATATATATAATTTGATAGATTTAG | 1256 |
| | | | | | |
| | | QY | 1261 | AAAAGTAGTTGAAAGTAGCGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGG | 1320 |
| | | DB | 1257 | AAAAGTAGTTGAAAGTAGCGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGG | 1316 |
| | | | | | |
| | | QY | 1321 | AGAAAGTTTGAATCTATAGAAATTCAGTGGGACATCCATAAAGGAAAGTGTATAGAAT | 1380 |
| | | DB | 1317 | AGAAAGTTTGAATCTATAGAAATTCAGTGGGACATCCATAAAGGAAAGTGTATAGAAT | 1376 |
| | | QY | 1381 | TAAATGGGACAGAGGAGTAAATACCTTTTATGATATATAAATTTTGTATTGTTGATTCA | 1440 |
| | | DB | 1377 | TAAATGGGACAGAGGAGTAAATACCTTTTATGATATATAAATTTTGTATTGTTGATTCA | 1436 |
| | | | | | |
| | | QY | 1441 | TAAGATTATAAATCTATGTTTAAATGATATAAATTTTAAAAAATAATATATATTAATTT | 1500 |

Query Match 99.0%; Score 2030.8; DB 3; Length 2048;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 2046; Conservative 0; Mismatches 2; Indels 4; Gaps 1;

QY 1 CATGTGTGCCCTACAGCACATAGGCGCTGTTGGTTGAGAGAGCAGAAAGCTGCTCTGA 60
 DB 1 CATGTGTGCCCTACAGCACATAGGCGCTGTTGGTTGAGAGAGCAGAAAGCTGCTCTGA 60

QY 61 CTTCCTCTCTTTTGACCTGTTGTATAAAGAGTAGAATATTTTAAAGCTGCGAA 120
 DB 61 CTTCCTCTCTTTTGACCTGTTGTATAAAGAGTAGAATATTTTAAAGCTGCGAA 120

QY 121 TACTAACTCTCTCTCAACCTTCCGCTCTCTTTTCCAAACACTTTTATTAACCTTTTACT 180
 DB 121 TACTAACTCTCTCTCAACCTTCCGCTCTCTTTTCCAAACACTTTTATTAACCTTTTACT 180

QY 181 TCTCATTTCTTACTCCACTTCTTGTCTATAAGCAAGAAATCACTTCTTTTAAGCTAACCCA 240
 DB 181 TCTCATTTCTTACTCCACTTCTTGTCTATAAGCAAGAAATCACTTCTTTTAAGCTAACCCA 240

QY 241 AAGCGCTCAATAAAGACATTCATAAATGTATCTTTCATTTTATAGGATAACAATACGT 300
 DB 241 AAGCGCTCAATAAAGACATTCATAAATGTATCTTTCATTTTATAGGATAACAATACGT 300

QY 301 GAACAGGGTTATTTTAAAGTGTCAACAAATCTAATAATTTTACCTGGCCGCTGAACA 360
 DB 301 GAACAGGGTTATTTTAAAGTGTCAACAAATCTAATAATTTTACCTGGCCGCTGAACA 360

Db 1437 TAAGATTATAAATCTATGTTTATAATGATAATAAATTTTAAAAAATAACTATATTAAT 1496
Qy 1501 CTGATTAGTCGATTACCGCCCTTTTATAATTTTCAATACCTGAGTATATGAATAAATCAG 1560
Db 1497 CTGATTAGTCGATTACCGCCCTTTTATAATTTTCAATACCTGAGTATATGAATAAATCAG 1556
Qy 1561 TTATCTGMAAGCAAAATAATATCTTTTGTAAACAGCGTTTCGGTCAAAATGGGAAGTTTCATG 1620
Db 1557 TTATCTGMAAGCAAAATAATATCTTTTGTAAACAGCGTTTCGGTCAAAATGGGAAGTTTCATG 1616
Qy 1621 TGTATTCAATAGTTTTTAATATAAAGTAAATTTTAAATTAATTTTGTGTTTTCAGA 1680
Db 1617 TGTATTCAATAGTTTTTAATATAAAGTAAATTTTAAATTAATTTTGTGTTTTCAGA 1676
Qy 1681 AATTTAAATAAATATTGAGCATGGGAAGTTTCACGGGCATCTATTGAGCGACCTAGACT 1740
Db 1677 AATTTAAATAAATATTGAGCATGGGAAGTTTCACGGGCATCTATTGAGCGACCTAGACT 1736
Qy 1741 GTTTGAAACATGATGTCGGGTGACATCTATGACCTTTTCAACTCAAACTAGTGAATAAT 1800
Db 1737 GTTTGAAACATGATGTCGGGTGACATCTATGACCTTTTCAACTCAAACTAGTGAATAAT 1796
Qy 1801 GCATTCTAGATAATACCTTTTCAAAATTTCAACAAACACAGCTTTTAACTTTTCTTCAACG 1860
Db 1797 GCATTCTAGATAATACCTTTTCAAAATTTCAACAAACACAGCTTTTAACTTTTCTTCAACG 1856
Qy 1861 GATTGGAATCCTTTTCAAACTTTTAAATAAATAAATAAATGCAATATTGTAATAATTATC 1920
Db 1857 GATTGGAATCCTTTTCAAACTTTTAAATAAATAAATAAATGCAATATTGTAATAATTATC 1916
Qy 1921 AACACCTCAACATGATGTTAGCGTACTATAAATAGTGCTCTTGGTGCTCTACTATCAT 1980
Db 1917 AACACCTCAACATGATGTTAGCGTACTATAAATAGTGCTCTTGGTGCTCTACTATCAT 1976
Qy 1981 CACATCAATCTTACACCAAAACCTTTGAGCTTTAATTTTCTACTATTCTCAGCAATAAC 2040
Db 1977 CACATCAATCTTACACCAAAACCTTTGAGCTTTAATTTTCTACTATTCTCAGCAATCAG 2036
Qy 2041 ATTCTAAATATC 2052
Db 2037 ATTCTAAAGATC 2048

RESULT 6
AAV15144
ID AAV15144 standard; DNA; 2042 BP.
XX
AC AAV15144;
XX
DT 02-JUL-1998 (first entry)
XX
DE New promoter used for root-specific expression in plants.
XX
KW Promoter; root; carrot; Kuroda Gosun; root-specific expression;
KW plant cell; soil pathogen; improve; nutritive value; edible root plant;
KW ss.
XX
OS Daucus carota.
XX
PN EP824150-A2.
XX
PD 18-FEB-1998.
XX
PF 12-AUG-1997; 97EP-00113923.
XX
PR 12-AUG-1996; 96JP-00212680.
XX
PA (SUMO) SUMITOMO CHEM CO LTD.
XX
PI Torikai S, Oeda K;
XX WPI; 1998-122310/12.
DR

XX
PT New carrot root gene, promoter and terminator - useful in genetic
PT engineering for directing root-specific gene expression.
XX
PS Claim 2; Page 15-16; 31pp; English.
XX
CC The present sequence represents a novel promoter, and is isolated from
CC the roots of carrots, cultivar Kuroda Gosun. The promoter can be used to
CC direct root-specific expression in plant cells. Since the promoter
CC enables expression of a desired protein in the roots of a plant, it is
CC useful in combat against pathogenic soil fungi and pests which are
CC difficult to kill by chemicals. It can also be used to improve the
CC nutritive value of edible root plants
XX
SQ Sequence 2042 BP; 719 A; 356 C; 323 G; 644 T; 0 U; 0 Other;
Query Match 11.7%; Score 239.6; DB 2; Length 2042;
Best Local Similarity 72.6%; Pred. No. 8.7e-35;
Matches 366; Conservative 0; Mismatches 129; Indels 9; Gaps 4;
Qy 1554 AATCAGTTTATCTGAAAGCAAAATAATATCTTTGTAAACACGCG---TTCGGTCAAATG 1609
Db 1543 AAATATTATTATCTGAATGATAACATCTTTGTAAACAAAACTGCGCAATAGGACCATAA 1602
Qy 1610 GGAAGTTCAATGTTTCAATAGTATTAAATAAAGTAAATTTTAAATTAATTTGTTATT 1669
Db 1603 CCNAGTTCAAGTATTCTTAAATGTTAATACTAACATGAGTATTTCCTT--TTCAAGGT 1660
Qy 1670 TTTGTTTCAGAAATTTAAATAAATATTGAGCATGGGAAGTTTCAGGGCATCATTTGAGC 1729
Db 1661 ATAAGTTAATCTTCAATCAATTAATTTTAAATTTGGACATATTGAGCAACTTTATGCC 1720
Qy 1730 AGCCTAGACTCTTTGAAACAATGATGTCGGGTGACATCTATGACCTTTCAACTCAAAC 1789
Db 1721 CACGTTGTTATTGTTTAAACAACGTTTGTCCGGTGATATTATGACCTTTCAACTCAAG 1780
Qy 1790 TAGTGAAT--AATGCAATTTAGAAATACATCTTTTCAAAATTTCAACAAACACAGCTTTAACT 1848
Db 1781 TAGCCAGTGAATGCTTTCTAGAAATATATCTTTTGAATTTTCAACAAACACAGCACTAACT 1840
Qy 1849 TTTCTTTCAACGGATTGGAATCCTTTTCAAACTTTTAAATTAATAAATAAATGCAATATT 1908
Db 1841 TTTCTTTTAAACAGATTAGAATCGTTTCGTAAACTTTTAAATTT--AAAAATACATTACT 1898
Qy 1909 GTAATATTATCAACACCTCAACATTTGATGTTAGCGTACTATAAATAGGTGCTCTTGGTG 1968
Db 1899 ATAATATTATTCAACACCTCAACATTTGATGTTAGCGTACTATAAATAGGTGCTCTTGGTG 1958
Qy 1969 CTCTACTATCATCATCAATCTTACACCAAAACCTTGAAGCTTAATTTTCTACTATT 2028
Db 1959 CTCTACTATCATCATCAATCTTCCAGCAAAACCTTGAAGCTTAATCTTCTACATAAT 2018
Qy 2029 CTCAGCAATAACATTCTAAATATC 2052
Db 2019 TTTAGCAAAACATCTTAAAGGTC 2042
RESULT 7
AAV15143
ID AAV15143 standard; DNA; 247 BP.
XX
AC AAV15143;
XX
DT 02-JUL-1998 (first entry)
XX
DE New promoter used for root-specific expression in plants.
XX
KW Promoter; root; carrot; Kuroda Gosun; root-specific expression;
KW plant cell; soil pathogen; improve; nutritive value; edible root plant;
KW ss.
XX
OS Daucus carota.
XX

```

PN EP824150-A2.
XX
PD 18-FEB-1998.
XX
PF 12-AUG-1997; 97EP-00113923.
XX
PR 12-AUG-1996; 96JP-00212680.
XX
PA (SUMO ) SUMITOMO CHEM CO LTD.
XX
XX Torikai S, Oeda K;
XX WPI; 1998-122310/12.
XX
PT New carrot root gene, promoter and terminator - useful in genetic
PT engineering for directing root-specific gene expression.
XX
PS Claim 1; Page 14; 31pp; English.
XX
CC The present sequence represents a novel promoter, and is isolated from
CC the roots of carrots, cultivar Kuroda Gosun. The promoter can be used to
CC direct root-specific expression in plant cells. Since the promoter
CC enables expression of a desired protein in the roots of a plant, it is
CC useful in combat against pathogenic soil fungi and pests which are
CC difficult to kill by chemicals. It can also be used to improve the
CC nutritive value of edible root plants
XX
SQ Sequence 247 BP; 88 A; 50 C; 21 G; 88 T; 0 U; 0 Other;

Query Match 9.6%; Score 196.6; DB 2; Length 247;
Best Local Similarity 89.6%; Pred. No. 5.8e-27;
Matches 223; Conservative 0; Mismatches 24; Indels 2; Gaps 1;

QY 1804 TTCTAGAATACATCTTTTCAAATTTCAACAAACACAGCTTTAACTTTCTTTCAACGGAT 1863
DB 1 TTCTAGAATATATCTTTTGAATTTCAACAAACACAGCAGCTTTCTTTTAAACAGAT 60

QY 1864 TGGAACTCTTTCTAAACTTTTAAATAAAATAAAATGCAATATATGTAATATTTATCAAC 1923
DB 61 TAGAATCGTTTCTCTAAACTTTTAAATTTTCAACAAACACAGCAGCTTTCTTTTAAACAGAT 118

QY 1924 ACCTCACATGTAGTTAGCGTACTATAAATAGGTGCTCTTGGTCTCTACTATCATCAC 1983
DB 119 ACCTCAACATTCATGTTAGCGTACTATAAATAGGTGCTCTTGGTCTCTACTATCATCAC 178

QY 1984 ATCAATCTTACCCACAAACCTTGAGCTTAATTTTCTACTTATTTCTCAGCAATAACATT 2043
DB 179 ATCAATCTTCCAGCAGCAAACTTGAGCTTAATCTTCTACTAATTTTTCAGCAAAACATT 238

QY 2044 CTAATAATC 2052
DB 239 CTAAGGTC 247

RESULT 8
AAZ49611
ID AAZ49611 standard; DNA; 246 BP.
XX
AC AAZ49611;
XX
XX 07-APR-2000 (first entry)
XX
XX Carrot CR16.3 fragment for synthesis of plant promoter.
XX
XX Synthetic DNA; plant promoter; CR16.3 fragment; carrot; transgenic plant;
XX soybean glycinin; stearyl-ACP-desaturase gene;
XX male sterility-related gene; ds.
XX
XX Daucus carota.
XX
XX EP976832-A2.
XX
XX 02-FEB-2000.
XX

XX 13-JUL-1999; 99EP-00113732.
XX
XX 15-JUL-1998; 98JP-00200372.
XX
XX (SUMO ) SUMITOMO CHEM CO LTD.
XX
XX Ishige F, Nishikawa S, Oeda K;
XX WPI; 2000-128374/12.
XX
XX Novel promoter used to produce transgenic plants with higher expression
XX of a desired gene.
XX
XX Claim 1; Page 11-12; 24pp; English.
XX
XX The present sequence is a CR16.3 fragment from carrot genomic DNA. It is
XX used for synthesis of a plant promoter which comprises nucleotides 112-
XX 246, 54-246, or 1-246 of this sequence and a synthetic DNA. The promoter
XX is used for controlling the expression of a desired gene e.g. soybean
XX glycinin, stearyl-ACP-desaturase and S-locus type specific RNase gene
XX (male sterility-related gene) in a host cell especially a microorganism
XX or a plant cell. The transformed plant cells can be used to produce
XX transgenic plants. The promoter is compact and therefore suitable for
XX higher expression of a desired gene in a particular tissue compared to
XX other host tissues
XX
SQ Sequence 246 BP; 88 A; 50 C; 21 G; 87 T; 0 U; 0 Other;

Query Match 9.5%; Score 195.6; DB 3; Length 246;
Best Local Similarity 89.5%; Pred. No. 8.9e-27;
Matches 222; Conservative 0; Mismatches 24; Indels 2; Gaps 1;

QY 1805 TCTAGAATACATCTTTTCAAATTTCAACAAACACAGCTTTAACTTTCTTTCAACGGATT 1864
DB 1 TCTAGAATATATCTTTTGAATTTTCAACAAACACAGCAGCTTTCTTTTAAACAGATT 60

QY 1865 GGAATCCTTTTCTAAACTTTTAAATAAAATAAAATGCAATATATGTAATATTTATCAACA 1924
DB 61 AGAATCGTTTCTCTAAACTTTTAAATTTTCAACAAACACAGCAGCTTTCTTTTAAACAGATT 118

QY 1925 CCTCACATGTAGTTAGCGTACTATAAATAGGTGCTCTTGGTCTCTACTATCATCAC 1984
DB 119 CCTCAACATTCATGTTAGCGTACTATAAATAGGTGCTCTTGGTCTCTACTATCATCAC 178

QY 1985 TCAATCTTACCCACAAACCTTGAGCTTAATTTTCTACTTATTTCTCAGCAATAACATT 2044
DB 179 TCAATCTTCCAGCAGCAAACTTGAGCTTAATCTTCTACTAATTTTTCAGCAAAACATT 238

QY 2045 TAAATATC 2052
DB 239 TAAAGGTC 246

RESULT 9
ADP07499
ID ADP07499 standard; DNA; 196 BP.
XX
AC ADP07499;
XX
XX 29-JUL-2004 (first entry)
XX
XX Carrot DNA.
XX
XX Carrot; gene; ds; expression inducing promoter;
XX transcription start point; RNA polymerase II.
XX
XX Daucus carota.
XX
XX JP2004135597-A.
XX
XX 13-MAY-2004.
XX

```


CC used as expression inducers. Furthermore, the present invention describes
CC DNA, vectors, transformants and the process by which to prepare
CC transformants. This polynucleotide sequence is DNA encoding a carrot
CC promoter (SeqID 3) of the invention.

XX SQ Sequence 2865 BP; 963 A; 461 C; 396 G; 1045 T; 0 U; 0 Other;

Query Match 7.9%; Score 162.4; DB 10; Length 2865;
Best Local Similarity 82.5%; Pred. No. 1.5e-20;
Matches 198; Conservative 0; Mismatches 41; Indels 1; Gaps 1;

Qy 22 AGGCGCTGTTGGTTGAGAGAGCAGAGCTGCTTCTGACTTCTTCTTTTGGACCTGT 81

Db 1716 AGGCGCTGTTGGTTTATGGAATCAGAGCTGCTTCTGACTTCTTCTTTT-ACCGT 1774

Qy 82 TTGTATAAGAGTAGAATAATTTTTTAAAGCTGCGAATACTTCTCTCTCAAC 141

Db 1775 TTGTATAAGAGTAGAATAATTTTTTAAAGCTGCGAATACTTCTCTCTCAAC 1834

Qy 142 TTCCGCTTCTTTTCCAAACACTTTTAACTTTTCTTCTTCTTCTTCTTCTTCT 201

Db 1835 TTCTGCTTCTTTTCCAAACACTTTTAACTTTTCTTCTTCTTCTTCTTCTTCT 1894

Qy 202 TTGCTATAAGCAAGAAATCACTTCTTTTAACTTAACTTAACTTAACTTAACT 261

Db 1895 TTACTATAAGCAAGAAATCACTTCTTTTAACTTAACTTAACTTAACTTAACT 1954

RESULT 12

AAZ49616/C
ID AAZ49616 standard; DNA; 140 BP.

XX AC AAZ49616;

XX DT 07-APR-2000 (first entry)

XX DE Oligonucleotide-4 for synthesis of CR16.1 fragment for plant promoter.
XX Synthetic DNA; plant promoter; CR16.1 fragment; carrot; transgenic plant;
XX soybean glycinin; oligonucleotide; stearyl-ACP-desaturase gene;
XX male sterility-related gene; ss.

XX OS Daucus carota.

XX OS Synthetic.

XX PN EP976832-A2.

XX PD 02-FEB-2000.

XX PF 13-JUL-1999; 99EP-00113732.

XX PR 15-JUL-1998; 98JP-00200372.

XX PA (SUMO) SUMITOMO CHEM CO LTD.

XX PI Ishige F, Nishikawa S, Oeda K;

XX WPI; 2000-128374/12.

XX PT Novel promoter used to produce transgenic plants with higher expression
XX of a desired gene.
XX PS Disclosure; Page 14; 24pp; English.

XX CC The present sequence is an oligonucleotide (- chain) used to prepare a
XX CR16.1 DNA fragment comprising nucleotides 112-246 of the carrot CR16.3
XX fragment. The CR16.1 DNA fragment is ligated to a 10 bp synthetic DNA for
XX construction of a plant promoter. The promoter is used for controlling
XX the expression of a desired gene e.g. soybean glycinin, stearyl-ACP-
XX desaturase and S-locus type specific RNase gene (male sterility-related
XX gene) in a host cell especially a microorganism or a plant cell. The
XX transformed plant cells can be used to produce transgenic plants. The
XX promoter is compact and therefore suitable for higher expression of a

CC desired gene in a particular tissue compared to other host tissues
XX SQ Sequence 140 BP; 45 A; 17 C; 34 G; 44 T; 0 U; 0 Other;

Query Match 5.8%; Score 119; DB 3; Length 140;

Best Local Similarity 92.6%; Pred. No. 1.1e-12;
Matches 125; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1918 ATCAACACCTCAACATTGATGTTAGCGTACTATATAAATAGGTGCTCTTTGGTCTCTACTAT 1977

Db 140 ATCAACACCTCAACATTGATGTTAGCGTACTATATAAATAGGTGCTCTTTGGTCTCTACTAT 81

Qy 1978 CATCATCATCAATCTTACACCAAAACCTTGAGCTTAATTTTCTACTTATTTCTCAGCAAT 2037

Db 80 CATCATCATCAATCTTACACCAAAACCTTGAGCTTAATTTTCTACTTATTTTCTCAGCAAT 21

Qy 2038 AACATTTCTAAATATC 2052

Db 20 AACATTTCTAAAGGTC 6

RESULT 13

AAZ49615
ID AAZ49615 standard; DNA; 140 BP.

XX AC AAZ49615;

XX DT 07-APR-2000 (first entry)

XX DE Oligonucleotide-3 for synthesis of CR16.1 fragment for plant promoter.
XX Synthetic DNA; plant promoter; CR16.1 fragment; carrot; transgenic plant;
XX soybean glycinin; oligonucleotide; stearyl-ACP-desaturase gene;
XX male sterility-related gene; ss.

XX OS Daucus carota.

XX OS Synthetic.

XX PN EP976832-A2.

XX PD 02-FEB-2000.

XX PF 13-JUL-1999; 99EP-00113732.

XX PR 15-JUL-1998; 98JP-00200372.

XX PA (SUMO) SUMITOMO CHEM CO LTD.

XX PI Ishige F, Nishikawa S, Oeda K;

XX WPI; 2000-128374/12.

XX PT Novel promoter used to produce transgenic plants with higher expression
XX of a desired gene.
XX PS Disclosure; Page 13; 24pp; English.

XX CC The present sequence is an oligonucleotide (+ chain) used to prepare a
XX CR16.1 DNA fragment comprising nucleotides 112-246 of the carrot CR16.3
XX fragment. The CR16.1 DNA fragment is ligated to a 10 bp synthetic DNA for
XX construction of a plant promoter. The promoter is used for controlling
XX the expression of a desired gene e.g. soybean glycinin, stearyl-ACP-
XX desaturase and S-locus type specific RNase gene (male sterility-related
XX gene) in a host cell especially a microorganism or a plant cell. The
XX transformed plant cells can be used to produce transgenic plants. The
XX promoter is compact and therefore suitable for higher expression of a
XX desired gene in a particular tissue compared to other host tissues

XX SQ Sequence 140 BP; 44 A; 34 C; 17 G; 45 T; 0 U; 0 Other;

Query Match 5.8%; Score 119; DB 3; Length 140;

Best Local Similarity 92.6%; Pred. No. 1.1e-12;

Matches 125; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

| | | | |
|-----------|---|---|------|
| Qy | 1918 | ATCAACACCTCAACATGATGTTAGCGTACTATAAATAGGTGCTCTTGGTCTCTACTAT | 1977 |
| Db | 5 | ATCAACACCTCAACATTCATGTTAGCGTACTATAAATAGGTGCTCTTGGTCTCTACTAT | 64 |
| Qy | 1978 | CATCATCATCAATCTTACACACACACACCTTGAGCTTAATTTTCTACTATTTCTCAGCAAT | 2037 |
| Db | 65 | CATCATCATCAATCTTCCAGCACACACCTTGAGCTTAATTTTCTACTAAATTTTCTAGCAA | 124 |
| Qy | 2038 | AACATCTTAAATATC | 2052 |
| Db | 125 | AACATCTTAAAGTC | 139 |
| RESULT 14 | | | |
| ID | ABZ10246 | | |
| XX | ABZ10246 | | |
| AC | ABZ10246 | | |
| XX | ABZ10246 | | |
| DT | 16-JAN-2003 | (first entry) | |
| XX | Haematopoietic cell proliferation disorder related DNA sequence #386. | | |
| XX | Human; haematopoietic cell proliferation disorder; cytostatic; | | |
| KW | gene therapy; lymphocytic leukaemia; acute myelogenous leukaemia; | | |
| KW | cytosine methylation state; gene; ds. | | |
| XX | | | |
| XX | Homo sapiens. | | |
| OS | | | |
| XX | WO20027272-A2. | | |
| PN | | | |
| XX | 03-OCT-2002. | | |
| PD | | | |
| XX | | | |
| PF | 26-MAR-2002; 2002WO-EP003401. | | |
| XX | | | |
| PR | 26-MAR-2001; 2001US-0278333P. | | |
| XX | | | |
| PA | (EPIG-) EPIGENOMICS AG. | | |
| XX | | | |
| PI | Berlin K, Braun A, Disler J, Guetig D, Howe A, Mueller J; | | |
| PI | Olek A, Piepenbrock C, Adorjan P, Grabs G, Lesche R, Leu E; | | |
| PI | Lewin A, Lipscher E, Maier S, Model F, Mueller V, Otto T, Pelet C; | | |
| PI | Schwope I, Ziebarth H; | | |
| XX | | | |
| XX | WFI; 2003-018942/01. | | |
| DR | | | |
| XX | | | |
| PT | Detecting and differentiating between hematopoietic cell proliferative | | |
| PT | disorders, comprises contacting a target nucleic acid with a reagent that | | |
| PT | distinguishes between methylated and non-methylated CpG dinucleotides. | | |
| XX | | | |
| PS | Claim 28; SEQ ID NO 386; 117pp; English. | | |
| XX | | | |
| CC | The present invention describes a method for detecting and | | |
| CC | differentiating between haematopoietic cell proliferative disorders | | |
| CC | associated with at least 1 gene and/or their regulatory regions in a | | |
| CC | subject. The method comprises contacting a target nucleic acid in a | | |
| CC | biological sample obtained from the subject with at least 1 reagent, | | |
| CC | which distinguishes between methylated and non-methylated CpG | | |
| CC | dinucleotides within the target nucleic acid. ABZ09861 to ABZ1118 | | |
| CC | represent specifically claimed nucleotide sequences from the present | | |
| CC | invention. Oligonucleotides from the present invention can be used; for | | |
| CC | differentiating between healthy haematopoietic cells and proliferative | | |
| CC | disorder haematopoietic cells; for differentiating between acute | | |
| CC | lymphocytic leukaemia and acute myelogenous leukaemia; as probes for | | |
| CC | determining the cytosine methylation state and/or single nucleotide | | |
| CC | polymorphisms (SNPs) of haematopoietic cell proliferation disorder | | |
| CC | related sequences and their complements; and as primers for the | | |
| CC | amplification of haematopoietic cell proliferation disorder related DNA | | |
| CC | sequences. The nucleotide sequences from the present invention can also | | |
| CC | be used for detecting a predisposition to, differentiation between | | |
| CC | subclasses, diagnosis, prognosis, treatment and/or monitoring of | | |
| CC | haematopoietic cell proliferative disorders. The present method enables a | | |

QY 1502 TGATTAGTCGATACCGCTTTTATATATTTTACAACTACTGAGTAATATGATTAATTCAGT 1561
 Db 2926 AATTTTATGAAATGTAATAATTTAAATGAATATTAATTTGTTTAAATAATAATAAAAAAATTA 2985
 QY 1562 TATCTGAAAGCAAAATATATCTTTTGTAAACACGCGTTCGGTCAAAATGGAAGTTCATGT 1621
 Db 2986 TTGAAATAATGATATATTTTAAATGATGATGTTTAAATTTTATAATATAAATTAATAA 3045
 QY 1622 GTATTCAAATGTTTAAATATAAAGTAATAATTTTAAATTAATTAATTTGTTTTCGAA 1681
 Db 3046 TAAATAAATGTTTATAATTCGTAATGAAATTTGAAATTAATTTGTTTATTTTATGTAATTT 3105
 QY 1682 ATTTAAATAAATATTTGAGCATGGGAGTTCAGGCGCATCATTTGACGACACTAGACTG 1741
 Db 3106 AATAAAAAATTAATTTTAAATGAATTTAAATGATGATGATGATGATGATGATGATGAT 3165
 QY 1742 ---TTTGAACAATGATGTCGGGTGATCATCTATGACCTTTCAACTCAAACTAGTGAATA 1798
 Db 3166 AATTTTAAATAAATATTTTAAATAAATTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTT 3225
 QY 1799 ATGCATTTCTAGATACATCTTTTCAAAATTTCAACAAACACAGCTTTTAACTTTTCTTCAA 1858
 Db 3226 AATTTATTAATTTTAAATAAATTTGATTAATTTTAAATTAATTAATTAATTAATTAAT 3285
 QY 1859 CGAATTCGAATCCTTTTCTAAACTTTTAAATAAATAAATAAATGATTAATTTGTAATTTA 1918
 Db 3286 AATTTTGTGTTATTTGTTGTTATTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTA 3345
 QY 1919 TCAACACCTCAACATTCATGTTA 1941
 Db 3346 AATGAATAAATTTGAAATGA 3368

RESULT 15

ID ABZ10100 standard; DNA; 8056 BP.
 XX AC ABZ10100;

DT 16-JAN-2003 (first entry)

XX Haematopoietic cell proliferation disorder related DNA sequence #240.

XX Human; haematopoietic cell proliferation disorder; cytostatic;
 KW gene therapy; lymphocytic leukaemia; acute myelogenous leukaemia;
 KW cytosine methylation state; gene; ds.

XX Homo sapiens.

OS WO200277272-A2.

PN 03-OCT-2002.

XX 26-MAR-2002; 2002WO-EP003401.

XX 26-MAR-2001; 2001US-0278333P.

XX (EPIG-) EPIGENOMICS AG.

XX Berlin K, Braun A, Distler J, Guetig D, Howe A, Mueller J;
 PI Olek A, Piepenbrock C, Adorjan P, Grabs G, Lesche R, Leu E;
 PI Lewin A, Lipscher E, Maier S, Model F, Mueller V, Otto T, Pelet C;
 PI Schwoppe I, Ziebarth H;

XX WPI; 2003-018942/01.

XX Detecting and differentiating between hematopoietic cell proliferative
 PT disorders, comprises contacting a target nucleic acid with a reagent that
 PT distinguishes between methylated and non-methylated CpG dinucleotides.

PS Claim 28; SEQ ID NO 240; 117pp; English.

XX The present invention describes a method for detecting and

CC differentiating between haematopoietic cell proliferative disorders
 CC associated with at least 1 gene and/or their regulatory regions in a
 CC subject. The method comprises contacting a target nucleic acid in a
 CC biological sample obtained from the subject with at least 1 reagent,
 CC which distinguishes between methylated and non-methylated CpG
 CC dinucleotides within the target nucleic acid. AB209861 to AB211118
 CC represent specifically claimed nucleotide sequences from the present
 CC invention. Oligonucleotides from the present invention can be used: for
 CC differentiating between healthy haematopoietic cells and proliferative
 CC disorder haematopoietic cells; for differentiating between acute
 CC lymphocytic leukaemia and acute myelogenous leukaemia; as probes for
 CC determining the cytosine methylation state and/or single nucleotide
 CC polymorphisms (SNPs) of haematopoietic cell proliferation disorder
 CC related sequences and their complements; and as primers for the
 CC amplification of haematopoietic cell proliferation disorder related DNA
 CC sequences. The nucleotide sequences from the present invention can also
 CC be used for detecting a predisposition to, differentiation between
 CC subclasses, diagnosis, prognosis, treatment and/or monitoring of
 CC haematopoietic cell proliferative disorders. The present method enables a
 CC highly specific classification of haematopoietic cell proliferative
 CC disorders allowing for improved and informed treatment of patients
 XX
 SQ Sequence 8056 BP; 3711 A; 371 C; 371 G; 3603 T; 0 U; 0 Other;

Query Match 4.5%; Score 91.8; DB 8; Length 8056;

Best Local Similarity 44.2%; Pred. No. 1.8e-07;

Matches 627; Conservative 0; Mismatches 767; Indels 25; Gaps 5;

QY 547 AATGCTTCTCAAAATGTTTTTATATGTAATAATAATGCCCATCAAGGATAAGTAAATTT 606

Db 1358 ATTTTATTAATAAATTTATTTTATATAAATAATATATCGATTTTTTTTATTTT 1417

QY 607 CCGGTTTAAACAGTTTGTATATATATATGTTTACACTTACAAGAGATATTCGTAATCT 666

Db 1418 TTTATTTTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1477

QY 667 TTTAGACGACAGAGACTTAGGTCAAAATGGCGCTGTAACACAGCTAGACTTGCTCA 726

Db 1478 ATAAATTAATTAATAATTAATATATATAAATAAATAAATAAATAAATAAATAAATAA 1537

QY 727 CTGATAAATAGATAATTTGTTAGTATATAATAGTAGGATCTACATGACATTAATAATAGA 786

Db 1538 TTAATAAATAATTTTAAAAAATAATTAATTAATAATTAATTAATTAATTAATTAATTA 1597

QY 787 GCTATTAAATTAAGTTACTATAATAAAGACAGGTTAGTAAACACAGAGAGTAAACA 846

Db 1598 TAATTTTAAATTTTAAAAAATCGAAATACGAATCGTAAATTAATAAATAAATAAATAA 1657

QY 847 AGAGCTTGCTGCTGTTAGTTAGTTAGTTAGTTAGTTAGTTAGTTAGTTAGTTAGTTAG 906

Db 1658 AATTAATAC-GATAAATAATTTTATTTTATTAATAAATAAATAAATAAATAAATAAATAA 1716

QY 907 GATCTAAGCACATAGAAATTTAGTACAGGTTAAACCTTTTACAAGAATTTATATTAAAC 966

Db 1717 TATATATTTATAATTTTAAATTTTAAATAAATAAATAAATAAATAAATAAATAAATAA 1776

QY 967 GAAAAATCAATTTTATAACATGCTCTCGGCTGTCTATTATAATAGGATCACTTACTGATCA 1026

Db 1777 AATTAATAAATAATTAAGTTATTTATTTTAAATAAATAAATAAATAAATAAATAAATAA 1836

QY 1027 TCCATTAAACCTGTTTAAACCAAAATCAATGAGATATAATATCTTCAATAGAAAGAG 1086

Db 1837 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1896

QY 1087 GACAATGCTCTTTGAAAAACAATAAGTAGTACTCCCTCCGCTCCCTCTGGAATGTATACAT 1146

Db 1897 AAAAAATAAATAAATTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1956

QY 1147 ATGGATTGGACACGAGACTTAAGAAAAATGATAAAGTAAATGTAGAGTAAAAAGAGAG 1206

Db 1957 AATTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2016

QY 1207 A-----AAGAAAAGTGGGTAAAGTAGCGGGGCCCAATATATATTAATTGATAGAT 1256

| | | | |
|----|------|--|------|
| Db | 2017 | ATATTTAAAAATAAATAATATATTTTAAATAATATTTTATTTTAAATAATTTTATATAAAAAATTTTATATAAAAAAT | 2076 |
| Qy | 1257 | TTAGAAAGTAGTTGAAAGTAGTGGGTGGGTGGGTATTTTATATATATAATAAAAAATTTTACTAT | 1316 |
| Db | 2077 | ATTTTAAATATATATAAATAATAATAATAATATTTTAATAATAATTTTAAATAATTTTTTTAT | 2136 |
| Qy | 1317 | TTTGAGAAAGTTTTGAAATGTATAGAATTCGATGGGACATCCATAAAAGGAAAGTGTATA | 1376 |
| Db | 2137 | TTTATTAAAAATAAANAATAAAAAAANAANAATAATAATTTTATTAAAAATTTAAAAAA | 2196 |
| Qy | 1377 | GAATTAATGGGACAGAGGGAGTAATACCTTTATGATATATAAATTTTTCGTATTTTGAT | 1436 |
| Db | 2197 | ACGAAAAAANAATAAAAAAANAANAATTTTAATTTTAATAAAAAATTTATTTTAAATA | 2256 |
| Qy | 1437 | TTCATATAGATTATAAATCTATGTATTATAATGATAATAATTTTAAAAAATAATACTATATT | 1496 |
| Db | 2257 | TTTAAAAATAAATTAANAANAATTTTAAATTTTAAATTTTATTTTAAAAAATAATTTATT | 2316 |
| Qy | 1497 | AAATCTGATTAGTCGATTTACCGCTTTTATAAATTTTACAATACTGAGTAATAATGAATAAA | 1556 |
| Db | 2317 | AAATTTTCGATTAAATAAATAATTTTATTTTAAATTAATTTTATTTTATTTTATTTTATAAA | 2376 |
| Qy | 1557 | TCAGTTATCTGAAAGCAAAAT-----AATATCTTTGTAAACAGCGCTTCGGTCAAAATGG | 1610 |
| Db | 2377 | TAAATTTATAAATTTTATTTATTTTCGAAAAATTAANAANAATAATTTTATTTATTTTAA | 2436 |
| Qy | 1611 | GAAGTTCATGTGTAATCAATAGTTTTTAATAATAAAGTAATTTTAAAT-----AATGTGT | 1666 |
| Db | 2437 | AAAATTAATTTTTTTTTTATTTTATTTATATATATTTTTTATTTTATTTTCGTTTTAAAAATAATT | 2496 |
| Qy | 1667 | ATTTTTCGTTTCAGAAATTTAAATAAATAATTTTGAGCATGGGAAGTTACCGGGCATCAATG | 1726 |
| Db | 2497 | AAAAATTAACGAAAAAATAAAAAAANAATAAATTAATTTATAAATAAATTTAAAAAATTTTA | 2556 |
| Qy | 1727 | AGCAGACTAGACTGTTTGAACATGTATGTCGGGTGTACATCTATGACCTTTCAACTCA | 1786 |
| Db | 2557 | AAAAATTTAAATTAATTTTTTAAAAAATAAANAATAAACAANAANAANAATTTATAAATATA | 2616 |
| Qy | 1787 | AACTAGTGAATAATGCATCTCAGAAATACATCTTTTCAAAATTTCAACAAACACAGCTTTAA | 1846 |
| Db | 2617 | ATAAATTTTAAAAAANAANAATAAANAANAAGTTTAAATTTTAAATTTTAAAAAATAATAA | 2676 |
| Qy | 1847 | CTTTTCTTCAACGGATTGGAATCCTTTTCTAAACTTTTTTAAAAATAAAAAAATGCAATTA | 1906 |
| Db | 2677 | AT----TTTTTAAATTTTTTAAATTTATTTTTTTTTTTTTTTTTTAAATAAAAAATAATTCGTTA | 2732 |
| Qy | 1907 | TTGTAATATTATCAACACCTCAACATTTGATTTAGCGT | 1945 |
| Db | 2733 | AAATAAATTTATTAATAAANAATAAANAANAATAATTTTAAATCGT | 2771 |

Search completed: March 14, 2005, 15:58:32
Job time : 1127 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 14, 2005, 15:39:49 ; Search time 355 Seconds
(without alignments)
9458.140 Million cell updates/sec

Title: US-09-806-197-1

Perfect score: 2052

Sequence: 1 catgtgtgcctacagcaca.....gcaataacattctaataatc 2052

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:*
1: /cgn2_6/ptodata/1/ina/5A COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|---------------------|
| 1 | 241.2 | 11.8 | 2042 | 2 | US-08-911-434A-2 |
| 2 | 196.6 | 9.6 | 247 | 2 | US-08-911-434A-1 |
| 3 | 195.6 | 9.5 | 246 | 3 | US-09-352-608-2 |
| 4 | 119 | 5.8 | 140 | 3 | US-09-352-608-6 |
| 5 | 119 | 5.8 | 140 | 3 | US-09-352-608-7 |
| 6 | 81.4 | 4.0 | 187169 | 4 | US-09-949-016-12776 |
| 7 | 81.4 | 4.0 | 181569 | 4 | US-09-949-016-15940 |
| 8 | 72.2 | 3.5 | 205044 | 4 | US-09-949-016-15851 |
| 9 | 72.2 | 3.5 | 205044 | 4 | US-09-949-016-15852 |
| 10 | 72.2 | 3.5 | 205044 | 4 | US-09-949-016-15853 |
| 11 | 72.2 | 3.5 | 223471 | 4 | US-09-949-016-12387 |
| 12 | 72.2 | 3.5 | 223471 | 4 | US-09-949-016-12724 |
| 13 | 72.2 | 3.5 | 223471 | 4 | US-09-949-016-12725 |
| 14 | 69.4 | 3.4 | 7218 | 1 | US-08-232-463-14 |
| 15 | 68.6 | 3.3 | 1141 | 4 | US-09-806-708B-22 |
| 16 | 68.6 | 3.3 | 18773 | 4 | US-09-949-016-14164 |
| 17 | 68.2 | 3.3 | 1141 | 4 | US-09-806-708B-22 |
| 18 | 67.8 | 3.3 | 205044 | 4 | US-09-949-016-15851 |
| 19 | 67.8 | 3.3 | 205044 | 4 | US-09-949-016-15852 |
| 20 | 67.8 | 3.3 | 205044 | 4 | US-09-949-016-15853 |
| 21 | 67.8 | 3.3 | 223471 | 4 | US-09-949-016-12387 |
| 22 | 67.8 | 3.3 | 223471 | 4 | US-09-949-016-12724 |
| 23 | 67.8 | 3.3 | 223471 | 4 | US-09-949-016-12725 |
| 24 | 67.2 | 3.3 | 601 | 4 | US-09-949-016-30530 |
| 25 | 67.2 | 3.3 | 601 | 4 | US-09-949-016-30531 |
| 26 | 67.2 | 3.3 | 601 | 4 | US-09-949-016-37149 |
| 27 | 67.2 | 3.3 | 601 | 4 | US-09-949-016-37150 |

Query Match 11.8%; Score 241.2; DB 2; Length 2042;
Best Local Similarity 72.8%; Pred. No. 5.4e-42;

ALIGNMENTS

RESULT 1

US-08-911-434A-2

; Sequence 2, Application US/08911434A

; Patent No. 5959176

; GENERAL INFORMATION:

; APPLICANT: TORIKAI, Satoshi

; APPLICANT: OEDA, Kenji

; TITLE OF INVENTION: PLANT PROMOTER AND UTILIZATION THEREOF

; NUMBER OF SEQUENCES: 14

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH, LLP

; STREET: P.O. BOX 747

; CITY: FALLS CHURCH

; STATE: VIRGINIA

; COUNTRY: UNITED STATES OF AMERICA

; ZIP: 22040

; COMPUTER READABLE FORM:

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/911,434A

; FILING DATE: 12-AUG-1997

; CLASSIFICATION: 800

; ATTORNEY/AGENT INFORMATION:

; NAME: Stewart, Raymond C.

; REGISTRATION NUMBER: 21,066

; REFERENCE/DOCKET NUMBER: 2185-0199P

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (703)205-8000

; TELEFAX: (703)205-8050

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2042 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; ORIGINAL SOURCE:

; ORGANISM: Daucus carota L.

; INDIVIDUAL ISOLATE: Kuroda Gosun

; FEATURE:

; NAME/KEY: promoter

; LOCATION: 1..2042

; US-08-911-434A-2

Sequence 37163, A
Sequence 37164, A
Sequence 145867, A
Sequence 145868, A
Sequence 146135, A
Sequence 146136, A
Sequence 146403, A
Sequence 146404, A
Sequence 30531, A
Sequence 37150, A
Sequence 37164, A
Sequence 145868, A
Sequence 146136, A
Sequence 146404, A
Sequence 156535, A
Sequence 16110, A
Sequence 11934, A
Sequence 13248, A

28 67.2 3.3 601 4 US-09-949-016-37163
29 67.2 3.3 601 4 US-09-949-016-37164
30 67.2 3.3 601 4 US-09-949-016-145867
31 67.2 3.3 601 4 US-09-949-016-145868
32 67.2 3.3 601 4 US-09-949-016-146135
33 67.2 3.3 601 4 US-09-949-016-146136
34 67.2 3.3 601 4 US-09-949-016-146403
35 67.2 3.3 601 4 US-09-949-016-146404
C 36 67 3.3 601 4 US-09-949-016-30531
C 37 67 3.3 601 4 US-09-949-016-37150
C 38 67 3.3 601 4 US-09-949-016-37164
C 39 67 3.3 601 4 US-09-949-016-145868
C 40 67 3.3 601 4 US-09-949-016-146136
C 41 67 3.3 601 4 US-09-949-016-146404
C 42 66.2 3.2 231129 4 US-09-949-016-156535
C 43 66.2 3.2 231129 4 US-09-949-016-16110
C 44 66.2 3.2 266293 4 US-09-949-016-11934
C 45 65.8 3.2 12313 4 US-09-949-016-13248

| Matches | 367; | Conservative | 0; | Mismatches | 128; | Indels | 9; | Gaps | 4; | |
|---------|------|--------------|-----------|------------|--------------|-------------|-------------|---------------|-----------------|------|
| Qy | 1554 | AAATCAGT | TATCTG | AAAGCAAA | TAATATCT | TGTAAACAGCG | ---- | TTCGGTCAAATG | 1609 | |
| Db | 1543 | AAATTTAT | CTCGAATG | ATAACATCT | TTGTAACAAAC | CTGGSCCAAT | ATAGACCATAA | 1602 | | |
| Qy | 1610 | GGAAGTTC | ATGTGTAAT | CAATAGT | TTTTTAATATA | AAAGTAAAT | TTTTAAATTA | TGTTATT | 1669 | |
| Db | 1603 | CCAAGTTC | ACGTGTAAT | CTAAAA | TGTTAACTAA | CACTAGTAT | TTTTCTT | --TTCAAAGGT | 1660 | |
| Qy | 1670 | TTTGTTTC | CAGNAAT | TTAAAA | TAAATTAAT | TGAGCAT | GGAAGTT | CACGGGCAT | CATTGACG | 1729 |
| Db | 1661 | ATAAGT | TTAAATCT | TTCAATCA | TTAATTAAC | TTTAAAT | TTGGACATTA | TATGAGCAAC | TTTATGCC | 1720 |
| Qy | 1730 | AGCACTAG | ACTGTT | TTGAAAC | AAATGAT | GTGCGGTGT | ACATCTAT | CAGACCT | TTTCAACTCAAAC | 1789 |
| Db | 1721 | CAGGTTGT | ATGTTTAA | CAACAG | TTTGTGCGGTGT | ATATTTAT | TATGACCT | TTTCAACTCAAGC | 1780 | |
| Qy | 1790 | TAGTGAAT | -AATG | CATTCT | TAGAAT | ACATCTTTT | CCAAATTT | CAACAAAC | CACAGCTTTAACT | 1848 |
| Db | 1781 | TAGCCAGT | GAAATG | CTTTCT | TAGAATAT | ATCTTTT | TGAAATTT | CAACAAAC | CACAGCACTAACT | 1840 |
| Qy | 1849 | TTTCTTT | CAACGGAT | CTGGAAT | CTCTTTT | CTCAAC | TTTTTTT | AAAAATA | AAAAATGCAATTATT | 1908 |
| Db | 1841 | TTTCTTT | TAACAGAT | TAGAAT | CGTTT | CTCTAACT | TTTTAAAA | TTT-- | AAAAAATACATTACT | 1898 |
| Qy | 1909 | GTAATAT | TTTATCA | CACCTCA | CAATCTG | ATGTTAG | CGGTACTATA | TAATAGGT | GCCTCTCGGTG | 1968 |
| Db | 1899 | ATAATAT | TTTATCA | CACCTCA | CAATCTG | ATGTTAG | CGGTACTATA | TAATAGGT | GCCTCTCGGTG | 1958 |
| Qy | 1969 | CTCTACT | ATCATCAT | CAATCTT | TACACCA | CAAACTTT | TGAGCTTAA | TTTTTTCT | TACTTTATT | 2028 |
| Db | 1959 | CTCTACT | ATCATCAT | CAATCTT | CCAGCACA | AACTTT | TGAGCTTAA | TCTTCT | TACTTAATT | 2018 |
| Qy | 2029 | CTCAGCA | ATAATCA | ATCTTA | ATAATC | 2052 | | | | |
| Db | 2019 | TTTAGCA | AAAAAC | ATCTTA | AAAGGTC | 2042 | | | | |

RESULT 2
US-08-911-434A-1
Sequence 1, Application US/08911434A
Patent No. 5959176
GENERAL INFORMATION:
APPLICANT: TORIKAI, Satomi
APPLICANT: OEDA, Kenji
TITLE OF INVENTION: PLANT PROMOTER AND UTILIZATION THEREOF
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH, LLP
STREET: P.O. BOX 747
CITY: FALLS CHURCH
STATE: VIRGINIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 22040
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/911,434A
FILING DATE: 12-AUG-1997
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Stewart, Raymond C.
REGISTRATION NUMBER: 21,066
REFERENCE/DOCKET NUMBER: 2185-0199P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)205-8000
TELEFAX: (703)205-8050
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:

```

; LENGTH: 247 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Daucus carota L.
; INDIVIDUAL ISOLATE: Kuroda Gosun
; FEATURE:
; NAME/KEY: promoter
; LOCATION: 1..247
; US-08-911-434A-1

Query Match          9.6%  Score 196.6;  DB 2;  Length 247;
Best Local Similarity 89.6%  Pred. No. 1e-32;
Matches 223;  Conservative 0;  Mismatches 24;  Indels 2;  Gaps 1;

Qy 1804  TTCTAGAATACATCTTTTTCAAATTTCAACAAACACAGCTTTTAACCTTTTCTTTCAACCGAT 1863
Db 1      TTCTAGAATATACTTTTGGAAATTTCAACAAACACAGCAGCTAACTTTCTCTTTTAAACAGAT 60

Qy 1864  TGGAAATCCCTTTTCTAAACCTTTTAAAAATAAAAAAATGCAATTATTGTAATATTTATCAAC 1923
Db 61     TAGAATCGTTTCTCAAACTTTTAAAAATT--AAAAAATACATTACTATAATTTATCAAC 118

Qy 1924  ACCTCAACATTGATGTTTAGGCGTACTATAAATAGTGTCTTTGGTGTCTCTACTATCATCAC 1983
Db 119    ACCTCAACATTGTTTAGGCTACTATAAATAGTGTCTTTGGTGTCTCTACTATCATCAC 178

Qy 1984  ATCAATCTTACACCACAAACCTTGAGCTTAATTTTTTCTACTTATTCTCAGCAATAACATT 2043
Db 179    ATCAATCTTCCAGCACAAACCTTGAGCTTAATCTTCTACTAATTTTTTAGCAAAACATT 238

Qy 2044  CTAATAATC 2052
Db 239   CTRAAGGTC 247

```

```

RESULT 3
US-09-352-608-2
; Sequence 2, Application US/09352608
; Patent No. 6218598
; GENERAL INFORMATION:
; APPLICANT: ISHIGE, Fumiharu
; APPLICANT: NISHIKAWA, Satomi
; APPLICANT: OEDA, Kenji
; TITLE OF INVENTION: Plant Promoter
; FILE REFERENCE: 2185-0353P
; CURRENT APPLICATION NUMBER: US/09/352,608
; CURRENT FILING DATE: 1999-07-13
; EARLIER APPLICATION NUMBER: 10-200372 JAPAN
; EARLIER FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 246
; TYPE: DNA
; ORGANISM: Daucus carota
US-09-352-608-2

Query Match          9.5%; Score 195.6; DB 3; Length 246;
Best Local Similarity 89.5%; Pred. No. 1.7e-32;
Matches 222; Conservative 0; Mismatches 24; Indels 2; Gaps 1;

Qy      1805  TCTAGATACATCTTTTCAATTTTCAACAACACAGCTTTTAACCTTTTCTTTTCAACGGATT 1864
Db      1      TCTAGAATATATCTTTTGAATTTTCAACAACACAGCACTAACTTTTCTTTTAAACAGATT 60

Qy      1865  GGAATCCTCTTTTCTAAACCTTTTAAAAATAAAAAAATGCATTTATTTGTAAATATTTATCAACA 1924
Db      61     AGAATCGTTTCCCTAAACCTTTTAAAAATT--AAAAAATACATTTACTATAATATTTATCAACA 118

```

```
Qy 1925 CCTCAACATTCATGTTAGCGTACTATAAATAGGTGCTCTTGGTCTCTACTATCATCACA 1984
Db 119 CCTCAACATTCATGTTAGCGTACTATAAATAGGTGCTCTTGGTCTCTACTATCATCACA 178
Qy 1985 TCAATCTTACACCAACAACTTGAGCTTAATTTTCTACTATTCTCTCAGCAATAACATTC 2044
Db 179 TCAATCTTCCAGCACAAACCTTGAGCTTAATCTTCTACTAAATTTTAGCAAAACATTC 238
Qy 2045 TAAATATC 2052
Db 239 TAAAGGTC 246

RESULT 4
US-09-352-608-6
; Sequence 6, Application US/09352608
; Patent No. 6218598
; GENERAL INFORMATION:
; APPLICANT: ISHIGE, Fumiharu
; APPLICANT: NISHIKAWA, Satomi
; APPLICANT: OEDA, Kenji
; TITLE OF INVENTION: Plant Promoter
; FILE REFERENCE: 2185-0353p
; CURRENT APPLICATION NUMBER: US/09/352,608
; CURRENT FILING DATE: 1999-07-13
; EARLIER APPLICATION NUMBER: 10-200372 JAPAN
; EARLIER FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 140
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:primer
US-09-352-608-6

Query Match 5.8%; Score 119; DB 3; Length 140;
Best Local Similarity 92.6%; Pred. No. 2.8e-16;
Matches 125; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1918 ATCAACACCTCAACATTCATGTTAGCGTACTATAAATAGGTGCTCTTGGTCTCTACTAT 1977
Db 5 ATCAACACCTCAACATTCATGTTAGCGTACTATAAATAGGTGCTCTTGGTCTCTACTAT 64
Qy 1978 CATCATCATCAATCTTACACCAACCTTGAGCTTAATTTTCTACTATTCTCAGCAAT 2037
Db 65 CATCATCATCAATCTTCCAGCACAAACCTTGAGCTTAATCTTCTACTAAATTTTAGCAAA 124
Qy 2038 AACATTCCTAAATATC 2052
Db 125 AACATTCCTAAAGGTC 139

RESULT 5
US-09-352-608-7/c
; Sequence 7, Application US/09352608
; Patent No. 6218598
; GENERAL INFORMATION:
; APPLICANT: ISHIGE, Fumiharu
; APPLICANT: NISHIKAWA, Satomi
; APPLICANT: OEDA, Kenji
; TITLE OF INVENTION: Plant Promoter
; FILE REFERENCE: 2185-0353p
; CURRENT APPLICATION NUMBER: US/09/352,608
; CURRENT FILING DATE: 1999-07-13
; EARLIER APPLICATION NUMBER: 10-200372 JAPAN
; EARLIER FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 140
; TYPE: DNA
```

```
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:primer
US-09-352-608-7

Query Match 5.8%; Score 119; DB 3; Length 140;
Best Local Similarity 92.6%; Pred. No. 2.8e-16;
Matches 125; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1918 ATCAACACCTCAACATTCATGTTAGCGTACTATAAATAGGTGCTCTTGGTCTCTACTAT 1977
Db 140 ATCAACACCTCAACATTCATGTTAGCGTACTATAAATAGGTGCTCTTGGTCTCTACTAT 81
Qy 1978 CATCATCATCAATCTTACACCAACCTTGAGCTTAATTTTCTACTATTCTCAGCAAT 2037
Db 80 CATCATCATCAATCTTCCAGCACAAACCTTGAGCTTAATCTTCTACTAAATTTTAGCAAA 21
Qy 2038 AACATTCCTAAATATC 2052
Db 20 AACATTCCTAAAGGTC 6

RESULT 6
US-09-949-016-12776
; Sequence 12776, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 12776
; LENGTH: 187169
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(187169)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12776

Query Match 4.0%; Score 81.4; DB 4; Length 187169;
Best Local Similarity 45.3%; Pred. No. 1.4e-07;
Matches 417; Conservative 0; Mismatches 496; Indels 8; Gaps 3;

Qy 883 ATTTCTTTAAAGTAATGTAAACTGATCTAAAGCACATAGAAAATTAGTACAGGTTAAAA 942
Db 465 ATATATATATAAAATATATATATATATAAATAAATAAATAAATAAATAAATAAATAA 524
Qy 943 CTTTTCACAGAAATTATTTAAACGAAATCATTTTATACATGCTCTCTCGGCTGTCATT 1002
Db 525 ATAAATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 584
Qy 1003 ATAAAGGATCACTTACTGATCATCCATTTAAACCTTGTAAACAAATTCATATGAGAT 1062
Db 585 TAAATATATATCTATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 644
Qy 1063 AAAATATCTTACAAATGAAAGAGGACAAATGTCTCTTTGAAAAAACAATAGTACTCCC 1122
Db 645 ATAAATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 704
Qy 1123 TCCGTCCCTCTGAAATGTATACATATGATGGACACGGACCTAAGAAAAATGTATAA 1182
```

```
Db 705 ATAATATATATAAAATATATAATATATATATATATAAAATATATAAAATATATAATAT 764
Qy 1183 GTAATGTAGAGTAAAAGAAAGAGAAAGTGGGTAAAGTAGCGGACCCACCAATA 1242
Db 765 ATATAAAATATAAAATATATATATATATATATATAAAATATATAAAATATATAAAATA 824
Qy 1243 TATAAATGTATAGATTTAGAAAAGTAGTTGAAAAGTAGTGGGTGGGTGGGATTTTATATTA 1302
Db 825 TATAATATATAAAATATATATATATATAAAATATATAAAATATATAAAATATATAAAATA 878
Qy 1303 TAAAAATTTACTATTTTGGAAAAGTTTGGAAAAGTTTGGAAAAGTTGAGTGGGACATCCATAA 1362
Db 879 TATAAAATATATAAAATATATAATATATAAAATATATAAAATATATAAAATATATAAAATAT 938
Qy 1363 AAGGAAGCTATAGACTGTTTGAACATGTATGTCGGGTGATACATCTATGACCTTTCAA 1782
Db 1297 ATAAATAAAATATAATATAAAATATAAAATATAAAATATAAAATATAAAATATAAAATA 1356
Qy 1783 CTCAAACTAGTGAATATGCA 1803
Db 1357 TATATTTAGTTCTAAGACA 1377

RESULT 7
US-09-949-016-15940
; Sequence 15940, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 15940
; LENGTH: 191569
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(191569)

; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15940
Query Match 4.0%; Score 81.4; DB 4; Length 191569;
Best Local Similarity 45.3%; Pred. No. 1.4e-07; Indels 8; Gaps 3;
Matches 417; Conservative 0; Mismatches 496;

Qy 883 ATTTCCTTTAAAAAGTAAATGTAAAGCTGATCTAAAGACACATAGAAAATTTAGTACAGGTTAAAA 942
Db 465 ATATATATATAAAATATATAATATATAAAATATATAAAATATATAAAATATATAAAATATAT 524
Qy 943 CTTTTCACAGAATTTATATATAAACGAAAATCAATTTTATAACATGCTCTCGGGCTGTCAAT 1002
Db 525 ATAATATATAAAATATATAATATATAAAATATATAAAATATATAAAATATATAAAATATATA 584
Qy 1003 ATAATAGGATCACCTTACTGATCAATCCATTAACCTTTTAAACCAATTCATAGAGAT 1062
Db 585 TAAATATATATATATATATAAAATATATAAAATATATAAAATATATAAAATATATAAAATAT 644
Qy 1063 AAAATATCTTACAATGAAAAGGAGCAATGCTCTTTGAAAAAACAATAAGTACTCCC 1122
Db 645 ATATATATATAAAATATATAAAATATATAAAATATATAAAATATATAAAATATATAAAATAT 704
Qy 1123 TCCGTCCCTCTGAAATGTATACATATGGATGGACACGGAGACTAAGAAAAATGTATAA 1182
Db 705 ATAATATATAAAATATATAAAATATATAAAATATATAAAATATATAAAATATATAAAATAT 764
Qy 1183 GTAATGTAGACTAAAAGAAAGAGAAAGTGGGTAAAGTAGCGGACCCACCAATA 1242
Db 765 ATATAAAATATAAAATATATATTTATATATAATATAAAATATATAAAATATATAAAATATA 824
Qy 1243 TATAATTTGATAGATTTTAGAAAAAGTAGTTGAAAAGTAGTGGGTGGGATTTTATATTA 1302
Db 825 TATAATATATAAAATATATAATATATAAAATATATAAAATATATAAAATATATAAAATATA 878
Qy 1303 TAAAAATTTACTATTTTGGAAAAGTTTGGAAAAGTTGATAGAAATTTGAGTGGGACATCCATAA 1362
Db 879 TATAAAATATATAAAATATATAATATATAAAATATATAAAATATATAAAATATATAAAATAT 938
Qy 1363 AAGGAAGCTATAGAAATTAATTTGAGGACAGAGGAGTAAATACCTTTTATCATATATAAAT 1422
Db 939 ATAAATATATAATATATAAAATATATATGGATATATAATATAATATATATATATAAAATA 997
Qy 1423 TTTGTTATTTTGTATTTTCAAGATTAATAATCTATGTTATAATGATAATAATATATATTTTAA 1482
Db 998 TATAATATATAAAATATAAAATATAAAATATAAAATATAAAATATAAAATATAAAATATA 1057
Qy 1483 AATAAATCTATATAATCTCGAATAGTCGATTAACCGCTTTTATAAATTTTACAATCTGA 1542
Db 1058 TAT-ATAATGTATAAAATATATAATATATAAAATATATAAAATATATAAAATATATA 1116
Qy 1543 GTAATATGAATTAATCAGTTATCTGAAAAGCAATAATATCTTTGTAAGACAGCGTTCGG 1602
Db 1117 ATATATAATATATAATATAATATATAATATAAAATATATAAAATATATAAAATATATA 1176
Qy 1603 TCAAACTAGTGAATATGCA 1803
Db 1357 TATATTTAGTTCTAAGACA 1377

RESULT 8
```

US-09-949-016-15851
; Sequence 15851, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15851
; LENGTH: 205044
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(205044)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15851

Query Match 3.5%; Score 72.2; DB 4; Length 205044;
Best Local Similarity 48.0%; Pred. No. 1.3e-05;
Matches 278; Conservative 0; Mismatches 288; Indels 13; Gaps 2;

Qy 1135 AAATGTATACATATGGATTGGACACGAGACTAAGAAAAATGTATAAGTAAGTACAGT 1194
Db 200520 ATATATTTATATATATTTTAAATATAATATAATAAATATATTTATATATTTAAAT 200579

Qy 1195 AAAAAGAAAGAGAAAGAGTGGGTAAAGTAGCGGGACCCACCAATATATAA----TTG 1250
Db 200580 ATAAATATATAAAATATATTTATATATATTTTAAATATAAATATAAATATATTTAT 200639

Qy 1251 ATAGATTTAGAAAGTAGTTGAAAGTAGTGGGTGGGATTTTATATATATAAAATTT 1310
Db 200640 ATATATTTTAAATATAAATATAATAATATATATTTTATATATATTTTAAATATAATATA 200699

Qy 1311 TACTATTTTGAGAAAGTTTTCGAAATGTATAGAATTGAGTGGGACATCCATAAAGGAAAG 1370
Db 200700 AATATATTTTATATATATTTTAAATATAAATATAAATATAAATATATTTTAAATATA 200759

Qy 1371 TGTATAGAATTAATGGGACAGAGGGAGTAACTCTTTATGATATATAAATTTTGTAT 1430
Db 200760 TATAAATATATAAATATATTTTATATATATTTTAAATATAAATATAAATATAAATATA 200819

Qy 1431 TTTGATTTTCATAGATTATATAAATCTATGTTATATGATATATAAATTTTAAATATAAATATA 1490
Db 200820 TTTTAAATATAAATATAAATATAAATATAAATATAAATATAAATATAAATATAAATATA 200879

Qy 1491 TATATTAATCTGATTAGTCGATTCGCCCTTTTATTAATTTTACAATCTAGTAATATG 1550
Db 200880 TATTTTAAATATAAATATAAATATAAATATAAATATAAATATAAATATAAATATAAATATA 200939

Qy 1551 AATAAATCAGTTATCTGAAAGCAAAATAATCTCTTTGTATAAAGCGCTTCGGTCAA---- 1606
Db 200940 TATATTTTAAATATAAATATAAATATAAATATAAATATAAATATAAATATAAATATAA 200999

Qy 1607 -----ATGGGAAGTTCATGTCATTCAATAGTTTAAATATAAAGTAAATTTAAATTA 1661
Db 201000 TATATATTTTAAATATAAATATAAATATAAATATAAATATAAATATAAATATAAATATA 201059

Qy 1662 TTGTTATTTTGTTCAGAAATTTTAAATATAAATTTATTTGA 1700
Db 201060 TATATATATTTTAAATATAAATATAAATATAAATATAAATATAAATATAAATATAAATATA 201098

RESULT 9

US-09-949-016-15852
; Sequence 15852, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15852
; LENGTH: 205044
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(205044)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15852

Query Match 3.5%; Score 72.2; DB 4; Length 205044;
Best Local Similarity 48.0%; Pred. No. 1.3e-05;
Matches 278; Conservative 0; Mismatches 288; Indels 13; Gaps 2;

Qy 1135 AAATGTATACATATGGATTGGACACGAGACTAAGAAAAATGTATAAGTAAGTACAGT 1194
Db 200520 ATATATTTATATATATTTTAAATATAAATATAAATATAAATATAAATATAAATATA 200579

Qy 1195 AAAAAGAAAGAGAAAGAGTGGGTAAAGTAGCGGGACCCACCAATATATAA----TTG 1250
Db 200580 ATAAATATATAAAATATATTTATATATATTTTAAATATAAATATAAATATAAATATA 200639

Qy 1251 ATAGATTTAGAAAGTAGTTGAAAGTAGTGGGTGGGATTTTATATATATAAAATTT 1310
Db 200640 ATATATTTTAAATATAAATATAATAATATATATTTTATATATATTTTAAATATAATATA 200699

Qy 1311 TACTATTTTGAGAAAGTTTTCGAAATGTATAGAATTGAGTGGGACATCCATAAAGGAAAG 1370
Db 200700 AATATATTTTATATATATTTTAAATATAAATATAAATATAAATATAAATATAAATATA 200759

Qy 1371 TGTATAGAATTAATGGGACAGAGGGAGTAACTCTTTATGATATATAAATTTTGTAT 1430
Db 200760 TATAAATATATAAATATATTTTATATATATTTTAAATATAAATATAAATATAAATATAA 200819

Qy 1431 TTTGATTTTCATAGATTATATAAATCTATGTTATATGATATATAAATTTTAAATATAAATATA 1490
Db 200820 TTTTAAATATAAATATAAATATAAATATAAATATAAATATAAATATAAATATAAATATAA 200879

Qy 1491 TATATTAATCTGATTAGTCGATTCGCCCTTTTATTAATTTTACAATCTAGTAATATG 1550
Db 200880 TATTTTAAATATAAATATAAATATAAATATAAATATAAATATAAATATAAATATAAATATA 200939

Qy 1551 AATAAATCAGTTATCTGAAAGCAAAATAATCTCTTTGTATAAAGCGCTTCGGTCAA---- 1606
Db 200940 TATATTTTAAATATAAATATAAATATAAATATAAATATAAATATAAATATAAATATAA 200999

Qy 1607 -----ATGGGAAGTTCATGTCATTCAATAGTTTAAATATAAAGTAAATTTAAATTA 1661
Db 201000 TATATATTTTAAATATAAATATAAATATAAATATAAATATAAATATAAATATAAATATAA 201059

Qy 1662 TTGTTATTTTGTTCAGAAATTTTAAATATAAATTTATTTGA 1700
Db 201060 TATATATATTTTAAATATAAATATAAATATAAATATAAATATAAATATAAATATAAATATA 201098

| Qy | 1662 | TTGTTATTTTGTTCAGAAATTTAAAAATAAAATTATTGA | 1700 |
|----|--------|--|--------|
| | | | |
| Db | 171487 | TATATATATTTTAAATATATAATAATATATAAAAAATATATTTA | 171525 |

RESULT 14
US-08-232-463-14/C

```

; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

| | | | | |
|-----------------------|-------------------|--------------------|-----------|--------------|
| Query Match | 3.4% | Score 69.4; | DB 1; | Length 7218; |
| Best Local Similarity | 4.6%; | Pred. NO. 2.4e-05; | | |
| Matches 16; | Conservative 212; | Mismatches 123; | Indels 0; | Gaps 0; |

QY 1051 ATTCAATGAGATAAATATCTTCAATGAAAGAAGCAATCTCTTTGAAAAACAA 1110

[illegible]

| | | | |
|----|------|--|------|
| Db | 1401 | RRR RRR RRR | 1344 |
| Qy | 1171 | AAAATGTATAAAGTAATGTAGAGTAAAAAGAAGAAGAAAGTGGTAAAGTAGCGG AAAATGTATAAAGTAATGTAGAGTAAAAAGAAGAAGAAAGTGGTAAAGTAGCGG AAAATGTATAAAGTAATGTAGAGTAAAAAGAAGAAGAAAGTGGTAAAGTAGCGG | 1230 |

| | Qy | Db | Qy | Db |
|------|--|------|----|----|
| 1291 | ATTTTATATTATAAAATTACTATTTTGGAGAAAGTTTGAATCATATAGATTGAGTG | 1350 | | |
| 1221 | RR | 1162 | | |
| 1351 | GGACATCCATAAAGAAAGTGTATAGAATTAAATCGGACAGAGGGAGTAA | 1401 | | |
| 1161 | RR | 1111 | | |

QY 1351 GGACATCCATAAAGGAAAGTGTATGAATTAAATGGGACAGAGGAGTAA 1401
::: : : ::::::::::: : : ::: : : ::::::::::: : :
Dp 1161 EEE 1111

```

RESULT 15
US-09-806-708B-22/c
; Sequence 22, Application US/09806708B
; Patent No. 6784342
; GENERAL INFORMATION:
; APPLICANT: The University of British Columbia
; TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
; FILE REFERENCE: 4810-58741
; CURRENT APPLICATION NUMBER: US/09/806,708B
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/147,133
; PRIOR FILING DATE: 1999-08-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
; LENGTH: 1141
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1)..(1141)
; OTHER INFORMATION: consensus sequence of A.t., L.a., and B.n. FAE1
US-09-806-708B-22

```

Query Match 3.3%; Score 68.6; DB 4; Length 1141;
Best Local Similarity 13.1%; Pred. No. 2.4e-05;
Matches 141; Conservative 392; Mismatches 536; Indels 10; Gaps 5;

QY 538 TTTTGGCCGAATGCTTCTCAAAATGTTTTTATATGTAAATAATGCCCATCCAAGGATA 597

| | | | |
|----|------|--|-----|
| DD | 1129 | TTKIKITTKANNNNNNNNMGKQDKNRMDAIKWSAIGIAJAWWINHAKRGAIMCMIWIIWJGIN | 107 |
| QY | 598 | AGTAAAAATCCCGTTTAAACCGATTGTGTAAATATATATATGTGTTCCTTCAACAGAGGATATT | 657 |

| | | | |
|----|------|--|-----|
| Db | 1069 | RRWCMRTYAMRTWYTRSNANWSCATKEMWWTMKWYATKYRTA-WYAMWCAWRNNMMWCAT | 101 |
| Qy | 658 | CGTAATACITTTTAGACGACAGAGACTTAGTCAAAAATGACGCTGGTAACAGCCTAG | 717 |

```

Db      1010  NGYKSCATNNMYYATTRWAAAYAAAKWARWAGNNMNYGAAGKNKGWCMAATWMBGW 951
          :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:|
Ov      718  ACTTGGTCACCTGATAAATAGATAAATGTTAGTATATAATAGTAGGATCTACAATGACATT 777

```

Db

950 ADTAGKCNNNNNWTTDVRMM--KAKNNNNNNAVYTACYNRAATNIGMATHWMMKWT 893

QY AAAATTAGACGCTATAATTAAGTACTAATAATAAGAAGGTGTATGAACAGAAAGCA 837
778 :
Db GAHSKRTRHHRTCRRTKYNNNNNNNARTVYYVHHAARWMAWTTNTNNNNNNNNNAC 833

[illegible]

QY 838 ATGTAACCTGATCTAAAGCACACATAGAAATTTAGTACAGTTTAAACTTTTCAAGAAATTT 957

D8 774 NNYMHAAVTTTTHCDWCYKTMNTWYDMMTMBTTTTTNNMTTSTNNNNNNNNNNWACTNN 715

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 14, 2005, 18:28:54 ; Search time 1188 Seconds
(without alignments)
10274.728 Million cell updates/sec

Title: US-09-806-197-1

Perfect score: 2052

Sequence: 1 catgtgacctacagcaca.....gcaataacattctaatatc 2052

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 5537552 seqs, 2974263231 residues

Total number of hits satisfying chosen parameters: 11075104

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
- 17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
- 18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
- 19: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
- 20: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
- 21: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
- 22: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|---------|----|----------------------|
| 1 | 99.4 | 4.8 | 8056 | 18 | US-10-473-126-386 |
| 2 | 93 | 4.5 | 3673778 | 16 | US-10-312-841-1 |
| 3 | 91.8 | 4.5 | 8056 | 18 | US-10-473-126-240 |
| 4 | 91.6 | 4.5 | 8056 | 18 | US-10-473-126-386 |
| 5 | 88.6 | 4.3 | 8056 | 18 | US-10-473-126-240 |
| 6 | 83 | 4.0 | 158001 | 17 | US-10-211-179-11 |
| 7 | 79.4 | 3.9 | 74665 | 18 | US-10-719-993-6854 |
| 8 | 75 | 3.7 | 11745 | 15 | US-10-240-453-206 |
| 9 | 71 | 3.5 | 5930 | 15 | US-10-311-455-490 |
| 10 | 70.6 | 3.4 | 3673778 | 16 | US-10-312-841-1 |
| 11 | 70 | 3.4 | 3252 | 13 | US-10-027-632-113786 |

| | | | | | | | |
|---|----|------|-----|---------|----|----------------------|-------------------|
| c | 12 | 70 | 3.4 | 3252 | 13 | US-10-027-632-113787 | Sequence 113787, |
| c | 13 | 70 | 3.4 | 3252 | 13 | US-10-027-632-113788 | Sequence 113788, |
| c | 14 | 70 | 3.4 | 3252 | 17 | US-10-027-632-113786 | Sequence 113786, |
| c | 15 | 70 | 3.4 | 3252 | 17 | US-10-027-632-113787 | Sequence 113787, |
| c | 16 | 70 | 3.4 | 3252 | 17 | US-10-027-632-113788 | Sequence 113788, |
| c | 17 | 69.8 | 3.4 | 6352 | 17 | US-10-221-613-195 | Sequence 195, App |
| c | 18 | 69.4 | 3.4 | 16258 | 17 | US-10-257-166-120 | Sequence 120, App |
| c | 19 | 69 | 3.4 | 3673778 | 16 | US-10-312-841-2 | Sequence 2, Appli |
| c | 20 | 68.6 | 3.3 | 49979 | 18 | US-10-741-601-5746 | Sequence 5746, Ap |
| c | 21 | 68.6 | 3.3 | 49979 | 19 | US-10-741-600-17905 | Sequence 17905, A |
| c | 22 | 68.2 | 3.3 | 15548 | 15 | US-10-311-455-2128 | Sequence 2128, Ap |
| c | 23 | 68 | 3.3 | 11422 | 15 | US-10-311-455-191 | Sequence 191, App |
| c | 24 | 68 | 3.3 | 11422 | 17 | US-10-257-166-17 | Sequence 17, Appl |
| c | 25 | 67.8 | 3.3 | 15732 | 14 | US-10-239-676-95 | Sequence 95, Appl |
| c | 26 | 67.8 | 3.3 | 15732 | 15 | US-10-240-453-107 | Sequence 107, App |
| c | 27 | 67.6 | 3.3 | 5984 | 18 | US-10-433-793-24 | Sequence 1984, Ap |
| c | 28 | 67.6 | 3.3 | 6050 | 15 | US-10-311-455-1984 | Sequence 1115, Ap |
| c | 29 | 67.4 | 3.3 | 5748 | 15 | US-10-311-455-1115 | Sequence 228, App |
| c | 30 | 67.4 | 3.3 | 18154 | 15 | US-10-311-455-228 | Sequence 1128, Ap |
| c | 31 | 67.2 | 3.3 | 5807 | 15 | US-10-311-455-1128 | Sequence 7676, Ap |
| c | 32 | 67.2 | 3.3 | 6816 | 18 | US-10-723-860-7676 | Sequence 97533, A |
| c | 33 | 67 | 3.3 | 2053 | 13 | US-10-027-632-97533 | Sequence 255, App |
| c | 34 | 67 | 3.3 | 2053 | 17 | US-10-027-632-97533 | Sequence 769, App |
| c | 35 | 67 | 3.3 | 115218 | 19 | US-10-278-698-255 | Sequence 11234, A |
| c | 36 | 67 | 3.3 | 115218 | 19 | US-10-278-698-769 | Sequence 394, App |
| c | 37 | 66.6 | 3.2 | 419 | 9 | US-09-960-352-11234 | Sequence 461, App |
| c | 38 | 66.6 | 3.2 | 6079 | 15 | US-10-311-455-394 | Sequence 1676, Ap |
| c | 39 | 66.6 | 3.2 | 6292 | 17 | US-10-221-714A-461 | Sequence 328, App |
| c | 40 | 66.4 | 3.2 | 18624 | 15 | US-10-311-455-1676 | Sequence 113786, |
| c | 41 | 66.2 | 3.2 | 1501 | 18 | US-10-473-126-328 | Sequence 113787, |
| c | 42 | 66.2 | 3.2 | 3252 | 13 | US-10-027-632-113786 | Sequence 113788, |
| c | 43 | 66.2 | 3.2 | 3252 | 13 | US-10-027-632-113787 | Sequence 113788, |
| c | 44 | 66.2 | 3.2 | 3252 | 13 | US-10-027-632-113788 | Sequence 113786, |
| c | 45 | 66.2 | 3.2 | 3252 | 17 | US-10-027-632-113786 | |

ALIGNMENTS

RESULT 1

US-10-473-126-386

; Sequence 386, Application US/10473126

; Publication No. US20040234973A1

; GENERAL INFORMATION:

; APPLICANT: Epigenomics AG

; TITLE OF INVENTION: Methods and nucleic acids for the analysis of hematopoietic cell

; TITLE OF INVENTION: proliferative disorders

; FILE REFERENCE:

; CURRENT APPLICATION NUMBER: US/10/473,126

; CURRENT FILING DATE: 2003-09-26

; NUMBER OF SEQ ID NOS: 1258

; SEQ ID NO 386

; LENGTH: 8056

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

US-10-473-126-386

Query Match 4.8%; Score 99.4; DB 18; Length 8056;
Best Local Similarity 44.5%; Pred. No. 1.1e-07;
Matches 624; Conservative 0; Mismatches 761; Indels 18; Gaps 5;

| | | | |
|----|------|---|------|
| Qy | 548 | ATGCTTCTCAAAATGTTTTTATATGTAATAATGATCCCAAGGATAAGTAAATTC | 607 |
| Db | 1975 | ATTAAATTAATTTTATTTTAAATTTTAAATTTTAAATTTTAAATTAATAT | 2034 |
| Qy | 608 | CCGTTAACCAAGTTGTTAATATATATATGTTACACTTACAGAGATATTCGTAATCTT | 667 |
| Db | 2035 | ATTTTAAATATTTTATTTTAAATAATTTTATTTAAATAATTTTAAATATATAAT | 2094 |
| Qy | 668 | TTAGACGACAGAGACTTAGTCAAAATCGACGCTGGTAAACAGCCTAGACTTGGTCAC | 727 |

2095 AATAAATTATAAATTATTTAATAAATTTTAAATAATTTTTTATTTTATTTTAAATAAATAA 2154
Qy 728 TGATAATAGATAATTCGTTAGTATATAATAGTAGGATCTACATGACATTTAAATTAGAG 787
Db 2155 TAAAAAATAAATAAATTTATTAATTTTAAATAATTTTAAATAAATGAAAAAATAAATAA 2214
Qy 788 CTATTAAATTAAGTTACTTAATAAATAAGAGAGGTTAGTAAACAGAAAGCAGGTAAAAACAA 847
Db 2215 AAAAAAATAAATTTTAAATAAATAAATTTTATTTTAAATAATTTAAATAAATAAATAA 2274
Qy 848 GAGCTTGCCTGCTGCTGTTAGTGTGTTGAGCTCAATTC-----TTTAAAGTAATGT 901
Db 2275 AATTTTAAATTTTAAATTTTATTTTAAATAAATTTTAAATAAATTTTATTTTAAATA 2334
Qy 902 AAAGTGNCTAAAGCAGATAGAAATTTAGTACAGGTTAAACCTTTTACAGAAATTTATAT 961
Db 2335 AATTTTATTTAAATTTTAAATAAATAATTTTATTTTATTTTAAATAAATTTATTTAAATTTAT 2394
Qy 962 TAAACGAAAAATCATTTTATAACATGCTCTCGGCTGTCATTAATAATAGGGATCACTTACT 1021
Db 2395 TATTTTGTGAAATTTAAATAAATAAATAATTTATTTTAAATAAATAATTTTAAATAATTTT 2451
Qy 1022 GATCATCCATTAAACCTGTTTAAACAAATTCATGAGATATAATCTTACAAATGAA 1081
Db 2452 TTATTTTATTTATATATTTTATTTTATTTTAAATAAATAAATAAATAAATAAATAAATAA 2511
Qy 1082 AGAAGGCAATGCTCTTTGAAATAAATAAATAAGTACTCCTCGGCTCCCTCGAATGTA 1141
Db 2512 AATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2571
Qy 1142 TACATATGGATTGGACACGGAGACTAAGAAAAATGTATAAGTAATGTAGAGTAAAAAGA 1201
Db 2572 TTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2631
Qy 1202 ARGAGAAAGAAAGTGGTAAAGTAGCGGACCCCAATATATAATTTGATAGATTAGA 1261
Db 2632 AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2691
Qy 1262 AAAGTAGTGAAGTAGTGGTGGGATTTTATTTATTTTAAATAAATAAATAAATAAATAAATAA 1321
Db 2692 TAAATTTATTTATTTTATTTTATTTTAAATAAATAAATAAATAAATAAATAAATAAATAA 2747
Qy 1322 GAAAGTTTGAATGTATAGAAATGAGTGGGACATCCATAAAGGAAAGTGTATAGAAAT 1381
Db 2748 AATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2807
Qy 1382 AAATGGGACAGAGGAGTAATACCTTTATGATATATAAATTTTGTGTTATTTGATTTTCAAT 1441
Db 2808 AAATTTGTTATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2865
Qy 1442 AAGATTATAAATCTATGTTATATGATAATATAAATTTTAAATAAATAAATAAATAAATAA 1501
Db 2866 TTTATGAAAAAATAATTTTGTGAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2925
Qy 1502 TGATTAGTCGATTACCGCTTTTATATTTTACAATCTAGTAAATATGATAAATAAATCAGT 1561
Db 2926 AATTTTATGAAATGTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2985
Qy 1562 TATCTGAAAAAGCAATAATATCTTTGTAACACAGCGTCCGGTCAAAATGGGAAGTTCATGT 1621
Db 2986 TTGAAAAATTTGATATATTTTAAATAATGATGATGTTTAAATTTTATATAAATAAATAAATAA 3045
Qy 1622 GTATTCATAGTTTATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1681
Db 3046 TAAAAATAAATGTTATATAATTTGGAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 3105
Qy 1682 ATTAAAAATAAATTTATGAGCATGGGAAGTTTCAGCGGCATTCATTGACGACACTAGACTG 1741
Db 3106 AATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 3165
Qy 1742 ---TTTGAACAATGATGTCGGGTGATACATCTATGACCTTTCAACTCAAACTAGTGAATA 1798
Db 3166 AATTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 3225

Qy 1799 ATGCATTTAGAAATACATCTTTTCAAATTTCAACAAACACAGCTTTTAACTTTTCTTTCAA 1858
Db 3226 AATTTATATATATTTTATTTTAAATAATTTGATTTATTTTAAATAATTTGATTTGATAAAA 3285
Qy 1859 CGGATTGGAATCCTTTTCTTAAACTTTTAAAAATAAATAAATAAATAAATAAATAAATAA 1918
Db 3286 AATTTTGTGTTATTTGTTGTTATTTTGTGAAAAAATAAATAAATAAATAAATAAATAAATAA 3345
Qy 1919 TCACACCTCAACATTTGATGTTA 1941
Db 3346 AAAATGAATAAAATTTGAAATGA 3368

RESULT 2
US-10-312-841-1/c
; Sequence 1, Application US/10312841
; Publication No. US20030186277A1
; GENERAL INFORMATION:
; APPLICANT: Epigenomics AG
; TITLE OF INVENTION: Diagnose von bedeutenden genetischen Parametern innerhalb des MHC
; FILE REFERENCE: E01/1208/WO
; CURRENT APPLICATION NUMBER: US/10/312,841
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 2
; SEQ ID NO 1
; LENGTH: 3673778
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; NAME/KEY: unsure
; LOCATION: (3294164)
US-10-312-841-1

Query Match 4.5%; Score 93; DB 16; Length 3673778;
Best Local Similarity 43.8%; Pred. No. 1.2e-05;
Matches 587; Conservative 0; Mismatches 745; Indels 7; Gaps 4;
Qy 566 TTTATATGTAAATAATAGCCCATCCAAAGGATAAGTAAATTTCCGTTTAAACCAAGTTTGT 625
Db 1714665 TATATATAAATAATATATATAAATAATATATATAAATAATATATAAATAATATATA 1714606
Qy 626 ATATATATGTTTACACTTTACAGAGGATATTCGTAATACTTTTACAGCAAGAGACTT 685
Db 1714605 TATATAAATAATATACATAAATAATATATATAAATAATATACATAAATAATATATA 1714546
Qy 686 AGTCAAAAATGGACGCTGGTAAACAGCCTAGACTTTGGTCACCTGATAAAATAGATAATTGT 745
Db 1714545 ATATATATATAAATAATATATAAATAATATATAAATAATATATAAATAATATATA 1714486
Qy 746 TAGTATAATATAGTAGGATCTCAATGACATTTAAATTTAGAGCTATTTAAATTAAGTTACTA 805
Db 1714485 AAAATATATAAATAATATATAAATAATATATAAATAATATATAAATAATATATAAATA 1714426
Qy 806 ATAAATAAGAGGTTAGTAAACAGAGCAAGCAGGTAAACCAAGCTTGCTGCTGCTGT 865
Db 1714425 ATATATAAATAATATATAAATAATATATAAATAATATATAAATAATATATAAATA 1714366
Qy 866 TTAGTTGTTGTGAGCTCATTTCTTTTAAAAAGTAAATGTAAACCTGATCTAAACACATAGAAA 925
Db 1714365 AAAATATATATAAATAATATATAAATAATATATAAATAATATATAAATAATATATA 1714306
Qy 926 TTTAGTACAGGTTAAAACTTTTACAAGAAATTTATATTTAAACGAAATCATTTTATAACAT 985
Db 1714305 ATATATAAATAATATATAAATAATATATAAATAATATATAAATAATATATAAATAAT 1714247
Qy 986 GTCTCTCGGCTGTCATTATAAATAGGGATCATTACTGATCATCCATTAAACCTTGTAA 1045
Db 1714246 ATATAAATAATATAAATAATATATAAATAATATATAAATAATATAAATAATATA 1714187
Qy 1046 AACAAATTCATGAGATAAATAATCT-TACAATGAAAAAGGACCAATGTCTCTTTGAAA 1104

US-10-473-126-240

| | | |
|-----------------------|------|--|
| Query Match | | 4.3%; Score 88.6; DB 18; Length 8056; |
| Best Local Similarity | | 45.4%; Pred. No. 9e-06; |
| Matches | | 442; Conservative 0; Mismatches 519; Indels 12; Gaps 3; |
| Qy | 736 | AGATAATCTTAGTATATATAGTAGGATCTACAAATGACATTAATAATAGAGCTATTAAT 795 |
| Db | 1399 | ATATTATTTTTATATAAAAAAATAAATTTTTTATATAAAAAATTAATAATTTAAATTTT 1340 |
| Qy | 796 | TAACTTACTAATAAATAGAGAGGTTAGTAAACAGAAAGCAGGTAATAAACAGAGCTTGC 855 |
| Db | 1339 | TTATTAATAATTAATAAATTTTTTTTATATATTTTTTATATTTTTTAAATAATATACGTTT 1280 |
| Qy | 856 | TGCTGTGTGTTAGTGTGTGAGCTCATTTCTTTAAAGTAATGTAAATGATCTAAAG 915 |
| Db | 1279 | TTATATATTAATAATAATACGTAACTGTTTTTTTTTTATTCGAAAAAATTTATTTATTTA 1220 |
| Qy | 916 | CACATAGAAATTTAGTACAGGTTAAACCTTTTACAGAAATTTATATTAACGAAATCAT 975 |
| Db | 1219 | AAAAAATTAATTAATAATTAATAATTTTAAATTAATTAATAATTAATAATCGAAATPATA 1160 |
| Qy | 976 | TTTATAACATGCTCTCGGCTGCTATTATATAGGATCACTTACTGATCATCCATTAAA 1035 |
| Db | 1159 | TTAATTTTAAATTTTCGATATCGAAAAAATTTTAAACGAAATTAATTAATCGATTTAA 1100 |
| Qy | 1036 | ACCTTGTAAACAAATTTCAATGAGATAAATATCTTTACAAATGAAAGAGGCAATGTC 1095 |
| Db | 1099 | AACGTTAAAAAAGAAACGAAATTAACGAAATTTTTTTTATTTACGTAATTTAAATTCGA 1040 |
| Qy | 1096 | TCCTTGAATAAACAATAGGTACTCCCTCGCTCGCTGAAATGTATACATATGATGG 1155 |
| Db | 1039 | AAATTAATAAATAATAAATAAATTTTTTAAATAATTTATTTTAAATAAATTAATAATA 980 |
| Qy | 1156 | ACAGGAGATTAAGAAAAATGATAAAGTAATGTAGAGTAAAGAGAAAGAGAAAG 1215 |
| Db | 979 | TTTTTATCGTTTAAATTAATAAATAAATAATTTATTTTAAATAAATTTATTAATAATA 920 |
| Qy | 1216 | TGGTAAAGTAGCGGAGCCACCAATATATAATTTGATAGATTTAGAAAAAGTAGTTGAAAG 1275 |
| Db | 919 | GACGAAATTAATTAATAATAAATTTTAAATTAATTTACGAAATTTTATACGTAAATAA 860 |
| Qy | 1276 | TAGTGGGTGGGGATTTTATATATAAATAATTTACTAT - TTTGAGAAAGTTTGA 1334 |
| Db | 859 | TAAAAAAGCGATTTTTTTTAAATAAATAATTTTAAATAATCGTTTATAAATTTATTTTA 800 |
| Qy | 1335 | TGTATAGAAATGTAGTGGGACATCCATAAAGGAAGTGTATAGATTTAAATGGGACAGAG 1394 |
| Db | 799 | AATATATAATTAATAATAATTTTATAAATTTTAAATTTTAAATTTTAAATAATTTAAT 740 |
| Qy | 1395 | GGAGTAATACCTTTATGATATATAAATTTTTTTTAT - - - - - TTTGATTTTCAAGATTA 1448 |
| Db | 739 | TATATATAAATTTTAAATTTTAAATTTTAAATTAATAAATAAATTTTTTTATATTTA 680 |
| Qy | 1449 | TAAATCTAGTATAATGATAATAAATTTTAAATAATACTATATATTAATTTCTGATAG 1508 |
| Db | 679 | TTTTTTTTTTTTTTTAAATAAATTTTATATTTTAAATAATTTTAAATAAATTTTAAAT 620 |
| Qy | 1509 | TCGATTAACCGCTTTTATAA - - - - - TTTTACATACCTGAGTAAATCAATAAATCACTTA 1563 |
| Db | 619 | TATATAATTTTTTTTTTAAATAATAATTTATTAATAATAAATAAATTTTAAATAAAGCTTA 560 |
| Qy | 1564 | TCGAAAGCAATAATAATCTTTTGAATAACAGCGCTCGGTCAATGGGAAGTTTCATGTGT 1623 |
| Db | 559 | TAATAATTAATAAATTTTAAATTTTAAATTAATAATAATTTTAAATTAATAATAATAA 500 |
| Qy | 1624 | ATCAATAGTTTAAATAAAGPAAATTTTAAATTAATTTGTTATTTTTTTTTCAGAAAT 1683 |
| Db | 499 | TTTTAAATTTTTTATTTAAATAATTTTATTTTAAATAATTTTATTTTATTTTATTTT 440 |
| Qy | 1684 | TTAAATAAATA 1696 |
| Db | 439 | TTAAAAAATAA 427 |

| | | |
|---|--------|---|
| RESULT 6 | | |
| US-10-211-179-11 | | |
| ; GENERAL INFORMATION: | | |
| ; APPLICANT: Nicholas M. Dean | | |
| ; APPLICANT: Kenneth W. Dobie | | |
| ; TITLE OF INVENTION: ANTISENSE MODULATION OF PHOSPHOTRYSYL PHOSPHATASE ACTIVATOR EXPRI | | |
| ; FILE REFERENCE: PUS-0011 | | |
| ; CURRENT APPLICATION NUMBER: US/10/211.179 | | |
| ; CURRENT FILING DATE: 2002-08-01 | | |
| ; NUMBER OF SEQ ID NOS: 119 | | |
| ; OTHER INFORMATION: n = a, t, c, or g | | |
| US-10-211-179-11 | | |
| Query Match 4.0%; Score 83; DB 17; Length 158001; | | |
| Best Local Similarity 46.0%; Pred. No. 0.00025; | | |
| Matches 458; Conservative 0; Mismatches 530; Indels 7; Gaps 5; | | |
| QY | 708 | ACAGCCTAGACTTGGTCACTGATATAAGATTAATTTGTTAGTATATAATAGTAGGATCTA 767 |
| Db | 126953 | AAATATATATATATATATAATAATAATATATATATATATATATATATATAATAATAATA 127012 |
| QY | 768 | CAATGACATTAATAATTAGAGCTATTAATTAAGTTTACTATAATAAAGAGAGGTTAGTAAA 827 |
| Db | 127013 | ATAATATATAATATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 127072 |
| QY | 828 | CAGAAAGCAGGTAAAAACAAGAGCTTGCTGCTGTGTTTGTGTTGTTGAGCTCATTTTC 887 |
| Db | 127073 | AAATAATATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 127132 |
| QY | 888 | TTTAAAGTAACTGAACTGATCTAAAGCAGATAGA - AATTTAGTACAGGTTAAAACTTT 946 |
| Db | 127133 | ATAATATATATATATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 127192 |
| QY | 947 | TACAAGATTTTATTAACGAAATCAATTTTATTAACATCTCTCGCTGCTGCTATTATAA 1006 |
| Db | 127193 | TAAATATATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 127252 |
| QY | 1007 | TAGGAGTCACTTACTGATCATCCATTTAAAAACCTTTTAAAAACCAATTCATAGATAAAA 1066 |
| Db | 127253 | AAATATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 127312 |
| QY | 1067 | TATCTTCAATG - AAAGAGGAGCAATGTCTCTTTGAAAAACAAATAGGACTCCCTCC 1125 |
| Db | 127313 | TATATAAATATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 127372 |
| QY | 1126 | GTCCCTCTGAATGTATACATATGGATTGGACAGGAGACTAAGAAAAATGTATAAAGTA 1185 |
| Db | 127373 | ATATTATATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 127432 |
| QY | 1186 | ATGTAGAGTAAAAAG 1245 |
| Db | 127433 | TAATATATATATATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 127492 |
| QY | 1246 | AATGTAGATTTAGAAAAAGTAGTGTGAAGTAGTGGGTGGGTGGGATTTTTTATATATAA 1305 |
| Db | 127493 | AAATATAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 127552 |
| QY | 1306 | AAATTTTACTATTTTGAGAAAGTTTGAATGTATAGAAATTTGAGTGGGACATCCATAAAG 1365 |
| Db | 127553 | AAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 127612 |
| QY | 1366 | GAAAGTGTATAGAAATTAATGGGACAGAGAGAGTA - ATACCTTTTATGATATATAAATTTT 1424 |
| Db | 127613 | TATATAAATAATAAAAAAATACATATATAATAATAATAATAATAATAATAATAATAATAATA 127672 |
| QY | 1425 | TGTTATTTTGTATTTCAAGATTTATAAATCT - ATGTTTATAATGATAATAATAATTTTAAAA 1483 |
| Db | 127673 | AAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 127732 |
| QY | 1484 | ATAAATACTATTAATTTCTGATGATGAGTACCGCTTTTATTAATTTTACAACTGAG 1543 |

;; PRIOR APPLICATION NUMBER: DE 10019173.8
;; PRIOR FILING DATE: 2000-04-07
;; PRIOR APPLICATION NUMBER: DE 10032529.7
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: DE 10043826.1
;; PRIOR FILING DATE: 2000-09-01
;; NUMBER OF SEQ ID NOS: 350
;; SEQ ID NO 206
;; LENGTH: 11745
;; TYPE: DNA
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
;; NAME/KEY: unsure
;; LOCATION: (9105)
US-10-240-453-206

Query Match 3.7%; Score 75; DB 15; Length 11745;
Best Local Similarity 42.7%; Pred. No. 0.0027;
Matches 499; Conservative 0; Mismatches 665; Indels 5; Gaps 2;

| | | | |
|----|------|--|------|
| Qy | 768 | CAATGACATTAATAATTAGAGCTATTAAATTAAGTTACTTAATAAATAAGAGAGGTTAGTAAA | 827 |
| Db | 4805 | CGAATAAATAATAATCTATTCTTAATAAATAAATAAATAAATAAATAAATAAATAAATAA | 4746 |
| Qy | 828 | CAGAAAGCAGGTAATAAACAAGAGCTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT | 887 |
| Db | 4745 | AATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA | 4686 |
| Qy | 888 | TTTAAAGTAGTAAGTAAGTCTTAAGCAGCATAGAAATTTAGTACAGGTTAAACCTTTT | 947 |
| Db | 4685 | CTCATTACACATCTACCTTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA | 4626 |
| Qy | 948 | ACAGAGATTTATATTAAGCAAGAAATCAATTTTATACATGCTCTCGGCTGTCATTATAT | 1007 |
| Db | 4625 | AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA | 4566 |
| Qy | 1008 | AGGATCACTTACTGATCATCCATTAATAAACCTTGTAAAAAATAAATAAATAAATAAATAA | 1067 |
| Db | 4565 | ACCTTAAATTTAACTCGAATCACCTTAAA---AAAAAACAATAAATAAATAAATAAATAA | 4509 |
| Qy | 1068 | ATCTTACAATGAAGAAGCAATGCTCTTTTGAAATAAATAAATAAATAAATAAATAAATAA | 1127 |
| Db | 4508 | ACCATAAACAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA | 4449 |
| Qy | 1128 | CCCTCTGAATGTATACATATGATTTGGACACGAGACTAAGAAATAATGTATAAGTAAT | 1187 |
| Db | 4448 | ACAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA | 4389 |
| Qy | 1188 | GTAGAGTAAAAAGAAAGAAAGAAAGTGGTAAAGTAGGGGACCCACCAATATATAA | 1247 |
| Db | 4388 | ATAACATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA | 4331 |
| Qy | 1248 | TTGTAGATTTAGAAAAAGTAGTTGAAAGTAGTGGGTGGGATTTTATATATATAAATAA | 1307 |
| Db | 4330 | TTTAAATCCTAAAAACAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA | 4271 |
| Qy | 1308 | ATTTACTATTTTGAAGAAGTTTGAATGTATAGATTTAGTGGGACATCCATAAAGGA | 1367 |
| Db | 4270 | TATTTTAAATTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA | 4211 |
| Qy | 1368 | AAGTGTATAGAAATTTAGGACAGAGGAGTAACTTTATCATATATAAATAAATAAATAA | 1427 |
| Db | 4210 | ATTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA | 4151 |
| Qy | 1428 | TATTTGATTTCAAGATTTATAAATCTATGTTATAATGATAAATAAATAAATAAATAAATAA | 1487 |
| Db | 4150 | TTCTACTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA | 4091 |
| Qy | 1488 | TACTATATTAATTTAGTTCGATTTACCGCCTTTTATATATATTTTAACTACTGAGTAAT | 1547 |
| Db | 4090 | TTTTCTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA | 4031 |

| | | | |
|----|------|--|------|
| Qy | 1548 | ATGAATAAATCAGTTATCTCGAAAGCAAAATAATATCTTTGTAAAAACAGCGTTCGGTCAAA | 1607 |
| Db | 4030 | AAAAAAACACAAATATTACAAAAACAAAAAATACTTACAAAAATAAAATCTTTAAAAA | 3971 |
| Qy | 1608 | TGGGAAGTTCATGTGTATTCATATAGTTTAAATAAAGTAAGTAAATTTTAAAAATTAATTTGTTA | 1667 |
| Db | 3970 | AAACAATTTAAAAATAAATAAATAAACAACAATAAATAAATAAATAAATAAATAAATAAATAA | 3911 |
| Qy | 1668 | TTTTTGTTCAGAAATTTAAAAATAAATTTATGAGCATGGGAAGTTTCAGGGGCATCATTTGA | 1727 |
| Db | 3910 | AAACTTTTCCATTTTACAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA | 3851 |
| Qy | 1728 | GCAGCACTAGACTGTTTGAACAATGTATGTCGGGTGTACATCTATAGACCTTTTCAACTCAA | 1787 |
| Db | 3850 | TAAAAATTAATAAATAAACAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA | 3791 |
| Qy | 1788 | ACTAGTGAATAATGCAATTTCTAGATATACATCTTTTCAAAATTTCAACAACAACAGCTTTAAC | 1847 |
| Db | 3790 | AACATAAACAAATTTCAATCAACCAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAA | 3731 |
| Qy | 1848 | TTTTCTTTCAACGATTTGGAATCCTTTTCTAAACCTTTTAAAAATAAATAAATAAATAAATAAATAA | 1907 |
| Db | 3730 | TTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA | 3671 |
| Qy | 1908 | TGTAATATTTTATCAACACCTCAACATTGA | 1936 |
| Db | 3670 | TTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA | 3642 |

RESULT 9

US-10-311-455-490/c

; Sequence 490, Application US/10311455

; Publication No. US20030143606A1

; GENERAL INFORMATION:

; APPLICANT: OLEK, Alexander

; APPLICANT: PIEPENBROCK, Christian

; APPLICANT: BERLIN, Kurt

; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determ

; FILE REFERENCE: 5013.1014

; CURRENT APPLICATION NUMBER: US/10/311.455

; CURRENT FILING DATE: 2002-12-16

; PRIOR APPLICATION NUMBER: PCT/EP01/07537

; PRIOR FILING DATE: 2001-07-02

; PRIOR APPLICATION NUMBER: DE 10032529.7

; PRIOR FILING DATE: 2000-06-30

; PRIOR APPLICATION NUMBER: DE 10043826.1

; PRIOR FILING DATE: 2000-09-01

; NUMBER OF SEQ ID NOS: 2424

; SEQ ID NO 490

; LENGTH: 5930

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

US-10-311-455-490

Query Match

Best Local Similarity 3.5%; Score 71; DB 15; Length 5930;

Matches 371; Conservative 0; Mismatches 420; Indels 8; Gaps 4;

| | | | |
|----|------|---|------|
| Qy | 897 | AATGTAACTGATCTAAGCAGACATAGAAATTTAGTACAGGTTAAACCTTTTACAGAAATT | 956 |
| Db | 5154 | ATTAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA | 5095 |
| Qy | 957 | TATATTAAACCAAAATCATTTTATAACATGCTCTCGGCTGTCTATTATATATAGGATCAC | 1016 |
| Db | 5094 | ATTTTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA | 5035 |
| Qy | 1017 | TTACTGATCATTCATTTAAACCTTTGTTAAACAATAATCAATGAGATATAATCTTACAA | 1076 |
| Db | 5034 | CTAATTTTATAATAATTTTCAAAACAATAATTTAAATAAATAAATAAATAAATAAATAAATAA | 4975 |


```

RESULT 11
US-10-027-632-113786/c
; Sequence 113786, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 113786
; LENGTH: 3252
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-113786

Query Match      3.4%; Score 70; DB 13; Length 3252;
Best Local Similarity 43.6%; Pred. No. 0.014;
Matches 474; Conservative 0; Mismatches 600; Indels 14; Gaps 3;

Qy      621 TTGTTATATATATGTTTACACTTACCAAGGAGGATATTCGTAATACTTTTAGACGACAAGA 680
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      2314 TTAATACTACAACATTTTATACCTTATTAATGATGTTTAATATGTTTATCTGTAATATA 2255
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy      681 GACCTTAGCTCAAAAATGGAGCGCTCGTAAACAGCCCTAGACACTTGGTCACTGATAAATAGATA 740

```

| | | | | | | | | | |
|---|------|---|------|--|--|--|--|--|--|
| US-10-027-632-113787/c | | | | | | | | | |
| ; Sequence 113787, Application US/10027632 | | | | | | | | | |
| ; Publication No. US20020198371A1 | | | | | | | | | |
| ; GENERAL INFORMATION: | | | | | | | | | |
| ; APPLICANT: Wang, David G. | | | | | | | | | |
| ; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide | | | | | | | | | |
| ; FILE REFERENCE: 108827.129 | | | | | | | | | |
| ; CURRENT APPLICATION NUMBER: US/10/027,632 | | | | | | | | | |
| ; CURRENT FILING DATE: 2002-04-30 | | | | | | | | | |
| ; PRIOR APPLICATION NUMBER: US 60/218,006 | | | | | | | | | |
| ; PRIOR FILING DATE: 2000-07-12 | | | | | | | | | |
| ; PRIOR APPLICATION NUMBER: US 60/198,676 | | | | | | | | | |
| ; PRIOR FILING DATE: 2000-04-20 | | | | | | | | | |
| ; PRIOR APPLICATION NUMBER: US 60/193,483 | | | | | | | | | |
| ; PRIOR FILING DATE: 2000-03-29 | | | | | | | | | |
| ; PRIOR APPLICATION NUMBER: US 60/185,218 | | | | | | | | | |
| ; PRIOR FILING DATE: 2000-02-24 | | | | | | | | | |
| ; PRIOR APPLICATION NUMBER: US 60/167,363 | | | | | | | | | |
| ; PRIOR FILING DATE: 1999-11-23 | | | | | | | | | |
| ; PRIOR APPLICATION NUMBER: US 60/156,358 | | | | | | | | | |
| ; PRIOR FILING DATE: 1999-09-28 | | | | | | | | | |
| ; PRIOR APPLICATION NUMBER: US 60/146,002 | | | | | | | | | |
| ; PRIOR FILING DATE: 1999-08-09 | | | | | | | | | |
| ; NUMBER OF SEQ ID NOS: 325720 | | | | | | | | | |
| ; SOFTWARE: FastSeq for Windows Version 4.0 | | | | | | | | | |
| ; SEQ ID NO 113787 | | | | | | | | | |
| ; LENGTH: 3252 | | | | | | | | | |
| ; TYPE: DNA | | | | | | | | | |
| ; ORGANISM: Human | | | | | | | | | |
| US-10-027-632-113787 | | | | | | | | | |
| Query Match | | | | | | | | | |
| Best Local Similarity 43.6%; Score 70; DB 13; Length 3252; | | | | | | | | | |
| Matches 474; Conservative 0; Mismatches 600; Indels 14; Gaps 3; | | | | | | | | | |
| Qy | 621 | TTGTTAATATATATGTTTACACTTACAAAGAGGATATTCGTAATACATTTTACAGACACAAGA | 680 | | | | | | |
| Db | 2314 | TTAATACTAACATTTTATACCTTATATTAATGAGTTAATATGTTTACTGTAATATA | 2255 | | | | | | |
| Qy | 681 | GACTTAGGTCANAAATGGAGCGCTGGTAAACAGCCTAGACTTGCTCACTGATAAATAGATA | 740 | | | | | | |
| Db | 2254 | TATTTAACTGCTATGTAACATATGAATGAAGTTAATATATATTAATTTAATATAATAAC | 2195 | | | | | | |
| Qy | 741 | ATTGTTAGTATATATAGTAGGATCTACATGACATTAATTTAGAGCTATTAAATAGT | 800 | | | | | | |
| Db | 2194 | ATTATAATATANAATTTTATAACAATATATCATNAATATTCATATTTATATATTTAATA | 2135 | | | | | | |
| Qy | 801 | TACTA-----ATAAATAAGAGAGGTTAGTAAACAGAAAGCAGGTAAACAAAG | 848 | | | | | | |
| Db | 2134 | TATAATACCATAATACTTATATTAATAAGCTAATTTAATATATTACATTAATCTAATA | 2075 | | | | | | |
| Qy | 849 | AGCTTGCTGCTGTTGTTAGTTGTTGTGAGCTCATTTCTTTAAAAAGTAATGTAACCTGA | 908 | | | | | | |
| Db | 2074 | ATCTTTTACTATTGTTGTAACCTTACATTAATTAATTAATATATTAATTTAACTTTATTA | 2015 | | | | | | |
| Qy | 909 | TCTAAAGCACATAGAATTTAGTACAGGTTA-AAACCTTTTACAAGAAATTTATATTAACG | 967 | | | | | | |
| Db | 2014 | TCTGGTTAACTATTATATATAGTTAAACCATATTTAAAAATATATAAATATATAAATAACA | 1955 | | | | | | |
| Qy | 968 | AAATCATTTTATAACATGCTCTCGGCTGCTCAATTAATAAAGGATCACTTACTGATCAT | 1027 | | | | | | |
| Db | 1954 | TGATATATTATATATAAATTTGCAATTTTATTAATTTATTTATATATATATTTAATAT | 1895 | | | | | | |
| Qy | 1028 | CCATTTAAACCTTTGTTAAACAAATTCATAGATAAAAATCTTACAAATGTAAGAAAGG | 1087 | | | | | | |
| Db | 1894 | ATAATATTTATATATATTTAATAAACAATTTTATATATATATATATATTTCAATAAAC | 1835 | | | | | | |
| Qy | 1088 | ACAATGCTCTTTGAAAAACAATAGGTACTCCCTCCGCTCGAATGTATACATA | 1147 | | | | | | |
| Db | 1834 | AATATTTATATATATATTTTAAATAACAATATATATATATATTTAATATACAATATT | 1775 | | | | | | |
| Qy | 1148 | TGGATTGGACACGAGACTAAGAAAAATGTATAAAGTAATGTAGAGTAAAAAGAAAGAGA | 1207 | | | | | | |

RESULT 13
US-10-027-632-113788/c
; Sequence 113788, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 113788
; LENGTH: 3252
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-113788

| | | | |
|----|------|--|------|
| Db | 1774 | TATATATTATATATATTTAATATACAAATATTTATATATATATATATATTTAATATACAATA | 1715 |
| Qy | 1208 | AGAAAGTGGGTAAAGTAGCGGACCCACCAATATATATTAATTCATAGATTTAGAAAAAGTA | 1267 |
| Db | 1714 | TTTATATATTTATATATATTTAATATACAATATTTATATATATATATATATTTAATATACA | 1655 |
| Qy | 1268 | GTTGAAAGTAGTGGGTGGGTGGGATTTTATATATATAAAAAATTTACTATTTTCAGAAAGT | 1327 |
| Db | 1654 | TATTTATATATATATATATTTAATATACAATATTTATATATATATATATATTTAATATATA | 1595 |
| Qy | 1328 | TTTGAATCTATAGAAATTCAGTGGGACATCCATAAAAAGGAAAGCTATAGAAATTAATGG | 1387 |
| Db | 1594 | TTTATATATTTATATATATTCGATGCATACATAAATTAATATCAATATATTTCAATAA | 1535 |
| Qy | 1388 | GACAGAGGAGTAATACCTTTATGATATATAAATTTTGTATTTTGAATTTTCATAAGATT | 1447 |
| Db | 1534 | TTAATATATGAATACCTATTAAATATATGAATATATTCCTATTCCTATATATTAATATAA | 1475 |
| Qy | 1448 | ATAAATCTATGTATTAATGATA-ATATAATTTTAAAAAATAACTATATATTAATCTGATT | 1506 |
| Db | 1474 | ATTATATATATATATTTGTTATATATATATATATATATAGATTTTAAATAAATCATTTGTA | 1415 |
| Qy | 1507 | AGTCGATTTACCGCTTTTATTAATTTTACAAATCTGAGTAATATGAATAAATCAGTTATCT | 1566 |
| Db | 1414 | TATATGTAATTCGATTTTATTAATGTTATATATAATGTAATGCAATATATATAATAA | 1355 |
| Qy | 1567 | GAAAACAAATATATCTTTTGTAAAAACAGCGTTCGGTCAAATGGGAAGTTTCATGTGATT | 1626 |
| Db | 1354 | TTATATAATATATATATATATATTTATTCATATATTTATATATATTTTAAATATGTTTAA | 1295 |
| Qy | 1627 | CAATAGTTTTAATATAAAGTAATTTTAAATTAATGTTATTTTGTTCAGAAATTTA | 1686 |
| Db | 1294 | TATAATAGTAGCATATTAATAAGTTAAATTTACTATATATTAATTAATTTAATATAAGTTAA | 1235 |
| Qy | 1687 | AAATAAAT 1694 | |
| Db | 1234 | CAGTAAGT 1227 | |

| Query Match | 3.4% | Score 70 | DB 13 | Length 252 |
|-----------------------|----------------|--|-----------|------------|
| Best Local Similarity | 43.6% | Pred. No. 0.014 | | |
| Matches 474 | Conservative 0 | Mismatches 600 | Indels 14 | Gaps 3 |
| Qy | 621 | TTGTTAAATATATATGTTTACACTTTACAGAGGATATTCGTAATACTTTTATAGACGACAAGA | 680 | |
| Db | 2314 | TTAAATCTAACAAATTTTATACCTTATATTAATGAGTTTAAATGCTTATATCTGTAATATA | 2255 | |
| Qy | 681 | GACTTAGGTCAAANAATGGACGCTGGTAAACAGCCCTAGACTTGGTCTACTGATAAATAGATA | 740 | |
| Db | 2254 | TATTTAACTGTTGTTAAACATATGAATGAATTAATTAATTAATTAATTAATTAATTAATTAAT | 2195 | |
| Qy | 741 | ATTGTTAGTATATAATATAGTAGGATCTACAATGACATTTAAAATTTAGAGCTATTAATTAAGT | 800 | |
| Db | 2194 | ATTATAAATAAATTTTATAACAAATATATCATTAATACATAATTCATATATATATTTTAATA | 2135 | |
| Qy | 801 | TACTA-----ATAAATAAGAGAGGTTAGTAAACAGAAAGCAGGTGAAAAACAAG | 848 | |
| Db | 2134 | TATTAATACCATATAACTTATTTTAAAGCTAAATTTTAATATATTAATTAATTAATTAATTAAT | 2075 | |
| Qy | 849 | AGCTTGCTGCTGTGTTTGTGTTAGTTGCTGAGCTCAATTTCTTTAAAAGTAATGTAAACTGA | 908 | |
| Db | 2074 | ATCTTTATACCTTATGTTTAACTTACATTAATAATTAATTAATTAATTAATTAATTAATTAAT | 2015 | |
| Qy | 909 | TCPTAAGCACAATAGAAATTTAGTACAGGTTA-AAACCTTTTACAAGAAATTTATATTAACG | 967 | |
| Db | 2014 | TCCTGTTAACTTATATATATATAGTTTAAACCATATTTAAAATATATAAATATATAAATAA | 1955 | |
| Qy | 968 | AAATCATTTTATAACATGCTCTCGGCTGTCATTAATAAGGATCACATTCAGTATCAT | 1027 | |
| Db | 1954 | TGATATATATATATATAAATTTGCACTTTTATTTATATATATATATATATATATATATAT | 1895 | |
| Qy | 1028 | CCATTAACAACTTGTGTTAAAAAATAATCCAATGAGATAAAAAATCTTCAATCAAAAGAAAGG | 1087 | |
| Db | 1894 | ATAATATTTATATATATATATATTTAAATPAAACAATATTTATATATATATATTTCAATPAA | 1835 | |
| Qy | 1088 | ACAATGCTCTTTGAAAAAACAATAGGTACTCCCTCCGTCCTGGAATGTATACATA | 1147 | |
| Db | 1834 | AAATTTATATATATATATATTTTAAATAACAATATATATATATATATATATATATATATAT | 1775 | |
| Qy | 1148 | TGGATTGGACGCGAGACTAAGAAAATGTATAAGTAAATGTAGAGTAAAGAAAGAGA | 1207 | |
| Db | 1774 | TATATATATATATATTTAAATATACAATATTTATATATATATATATATATATATATATAT | 1715 | |
| Qy | 1208 | AAGAAAAAGTGGTAAAGTAGCGGACCCCAATATATATAATTTGATGATTTAGAAAAAGTA | 1267 | |
| Db | 1714 | TTTATATATATATATATTTTAAATATACAATTTTATATATATATATATATATATATATAT | 1655 | |
| Qy | 1268 | GTTGAAAGTAGTGGGTGGGTGGGATTTTATATTAATAAAATTTTACTTTTGGAGAAAGT | 1327 | |
| Db | 1654 | TATTTATATATATATATTTTAAATATACAATTTTATATATATATATATATATATATATAT | 1595 | |
| Qy | 1328 | TTTGAAATGTATAGAAATTTGAGTGGGACATCCATPAAAAAGGAAGGTATAGAAATTAATGG | 1387 | |
| Db | 1594 | TTTATATATATATATAAATTTGATGTCATACATAAATTTATATATCAATATATTTTCATA | 1535 | |
| Qy | 1388 | GACAGGGAGTAAATACCTTTATGATATATAAAATTTTGTGTTTATTTTTCATTTTCATAAG | 1447 | |
| Db | 1534 | TTAATATATGAAATCTATTAATATATGAAATATATGCTATTTCTTATATTTAATATATAAGA | 1475 | |
| Qy | 1448 | ATAAATCTATGTTATAATATGATA-ATAAATTTTAAAAAATAACTATATATTAATTTCTGAT | 1506 | |
| Db | 1474 | ATTATATATATATATGTTAT | 1415 | |
| Qy | 1507 | AGTCGATTTACCGCTTTTATAAATTTTACAATACGTAGTAAATAGAAATCAAGTCACT | 1566 | |
| Db | 1414 | TATTTATGTAATTTGATTTTATGTTGTTATAAATAATGTAATGCAAAATATATATAAATATA | 1355 | |
| Qy | 1567 | GAAAGCAAAATAATCTTTTGTGTAACACAGCGTTCCGGTCAATGGGAGTTTCATGCTGAT | 1626 | |
| Db | 1354 | TTATATAATATATAATATATATATATTTTCAATATATTTATATATATTTTAAATGTTTAACT | 1295 | |
| Qy | 1627 | CAATAGTTTTAAATATAAAAGTAAATTTTAAATTTAAATTTGTTATTTTCTTTGTTTCA | 1686 | |

```

Db      1294 TATAATGTTAGCATATTAAATAAGTAAATTTACTATTAAATTAATTAATAAGTTAA 12335
        |||
Qy      1687 AAATAAAT 1694
        |||||
Db      1234 CAGTAAGT 1227
        |||

RESULT 14
US-10-027-632-113786/c
; Sequence 113786, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;   POLYMORPHISMS IN THE HUMAN GENOME
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 113786
; LENGTH: 3252
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-113786

```


| | | | |
|----|------|--|------|
| Qy | 1507 | AGTCGATTACCGCCTTTTATTAATTTTACAATACTGAGTAATATGCAATAAATCAGTTATCT | 1566 |
| Db | 1414 | TATTATGTAATTGATTTTATTAATGTTATATAATAATGTAATGCAAAATTATATATAATATA | 1355 |
| Qy | 1567 | GAAAAGCAATAATATCTTTGTAAAACAGCGTTCGGTCAATGGGAAGTTCATGTGTATT | 1626 |
| Db | 1354 | TTATATAATATAATATATATATTATTCATATATATATATTTTAATATGTTTAACTATA | 1295 |
| Qy | 1627 | CAATAGTTTAAATATAAAAAGTAAATTTTAAATTAATTGTTATTTTGTTCAGAAATTTA | 1686 |
| Db | 1294 | TATATAGTTAGCATATTTAATAGTTAATTACTATATTAATTAAATTTAATATAAGTTAA | 1235 |
| Qy | 1687 | AAATAAAT | 1694 |
| Db | 1234 | CAGTAAGT | 1227 |

Search completed: March 14, 2005, 23:09:30
Job time : 1198 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 14, 2005, 15:29:49 ; Search time 6848 Seconds
(without alignments)
11405.941 Million cell updates/sec

Title: US-09-806-197-1
Perfect score: 2052
Sequence: 1 catgtgtgcctacagcaca.....gcaataacattcttaataatc 2052

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST.*

1: gb_est1.*
2: gb_est2.*
3: gb_hic.*
4: gb_est3.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_gesi.*
9: gb_ges2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID | Description |
|------------|-------|---------------|--------|-------|--------------------|
| C 1 | 97.8 | 4.8 | 1101 | 9 | CNS00EVL |
| C 2 | 97.2 | 4.7 | 1101 | 9 | CNS0039G |
| C 3 | 95.6 | 4.7 | 1542 | 9 | AG386981 Mus muscu |
| C 4 | 95 | 4.6 | 1101 | 9 | CNS0039G |
| C 5 | 94.6 | 4.6 | 1758 | 9 | CU509408 SAIL_811 |
| C 6 | 93.6 | 4.6 | 1896 | 9 | CG753083 |
| C 7 | 93 | 4.5 | 1202 | 8 | CC262481 CH261-167 |
| C 8 | 92.4 | 4.5 | 1101 | 9 | AL061936 Drosophil |
| C 9 | 90.8 | 4.4 | 1101 | 9 | CNS0021J |
| C 10 | 90.4 | 4.4 | 1608 | 9 | CU118721 ISB1-72J8 |
| C 11 | 88.8 | 4.3 | 1067 | 6 | CD386564 |
| C 12 | 88.8 | 4.3 | 1268 | 9 | AG347098 Mus muscu |
| C 13 | 88.8 | 4.3 | 1392 | 9 | CG757503 P052-4-CO |
| C 14 | 88 | 4.3 | 1277 | 8 | CC253231 CH261-180 |
| C 15 | 88 | 4.3 | 1696 | 9 | AG346840 Mus muscu |
| C 16 | 87.6 | 4.3 | 1275 | 9 | CU033318 CH216-36F |
| C 17 | 87.2 | 4.2 | 887 | 9 | AG526041 Mus muscu |
| C 18 | 87 | 4.2 | 1275 | 9 | CU033318 CH216-36F |
| C 19 | 87 | 4.2 | 2087 | 9 | AG333887 Mus muscu |
| C 20 | 86.8 | 4.2 | 1348 | 9 | CG749499 P043-4-AO |
| C 21 | 85.6 | 4.2 | 1350 | 9 | CU019486 CH216-5G1 |
| C 22 | 85.4 | 4.2 | 1780 | 9 | AG320553 Mus muscu |
| C 23 | 85.2 | 4.2 | 1320 | 9 | CU103881 ISB1-42C8 |
| C 24 | 85 | 4.1 | 1539 | 9 | AG340947 Mus muscu |

| | | | | | |
|------|------|-----|------|---|--------------------|
| C 25 | 84.8 | 4.1 | 1254 | 9 | AG349719 |
| C 26 | 84.6 | 4.1 | 987 | 9 | CNS014PQ |
| C 27 | 84.6 | 4.1 | 1364 | 9 | CG757970 P053-2-CO |
| C 28 | 84 | 4.1 | 1101 | 9 | CNS017KE |
| C 29 | 83.8 | 4.1 | 822 | 7 | CK416977 AUF IpInt |
| C 30 | 83.6 | 4.1 | 1489 | 9 | AG350139 Mus muscu |
| C 31 | 83.4 | 4.1 | 1101 | 9 | CNS00FMC |
| C 32 | 83 | 4.0 | 1211 | 9 | AG349657 Mus muscu |
| C 33 | 83 | 4.0 | 1592 | 9 | CG750135 P044-3-D0 |
| C 34 | 82.8 | 4.0 | 1210 | 9 | CG749728 P044-1-C0 |
| C 35 | 82.6 | 4.0 | 1506 | 9 | AG278469 Mus muscu |
| C 36 | 81.6 | 4.0 | 1092 | 9 | CNS020K7 |
| C 37 | 81.6 | 4.0 | 1101 | 9 | CNS00EO7 |
| C 38 | 81.6 | 4.0 | 1355 | 9 | AG346348 |
| C 39 | 81.6 | 4.0 | 1745 | 9 | AG338221 Mus muscu |
| C 40 | 81.4 | 4.0 | 812 | 8 | BH178455 O11_J_02- |
| C 41 | 81.4 | 4.0 | 812 | 8 | AL615412 T3 end of |
| C 42 | 81.4 | 4.0 | 1050 | 9 | CNS013NS |
| C 43 | 81.4 | 4.0 | 1391 | 9 | CG754863 P050-2-G0 |
| C 44 | 81.2 | 4.0 | 1094 | 9 | CNS012F2 |
| C 45 | 81.2 | 4.0 | 1715 | 9 | AG288305 Mus muscu |

ALIGNMENTS

RESULT 1
LOCUS CNS00EVL/c 1101 bp DNA linear GSS 04-JUN-1999
DEFINITION Drosophila melanogaster genome survey sequence T7 end of BAC:
BACR29B23 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.

ACCESSION AL069706
VERSION AL069706.1 GI:4949849

KEYWORDS GSS.
SOURCE Drosophila melanogaster (fruit fly)

ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.

REFERENCE 1 (bases 1 to 1101)
Genoscope.

AUTHORS Direct Submission

TITLE Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :

JOURNAL BP 191 91006 EVRY cedex - FRANCE (E-mail : seque@genoscope.cns.fr
- web : www.genoscope.cns.fr)

COMMENT Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mamoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's pi and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES
Location/Qualifiers
1..1101
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR29B23"
/clone_lib="RPCI-98"
/notes="end : T7"

ORIGIN

Query Match 4.8%; Score 97.8; DB 9; Length 1101;
Best Local Similarity 34.8%; Pred. No. 8.7e-08;

Hattori, M., Toyoda, A., Noguchi, H., Kojima, T. and Sakaki, Y.
BAC end Sequences of Library MSMG01
Unpublished
2 (bases 1 to 1542)
Hattori, M., Toyoda, A., Noguchi, H., Kojima, T. and Sakaki, Y.
Direct Submission
Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail: hattori@gsc.riken.jp, URL: http://hgp.gsc.riken.go.jp/
Tel: 81-45-503-9111, Fax: 81-45-503-9170)
Clones are derived from the mouse BAC library MSMG01. For BAC
library availability, please contact Kuniya Abe (abe@crc.riken.jp).
Tsukuba Institute of Physical and Chemical Research (RIKEN) 3-1-1
Koyadai, Tsukuba, 305-0074 Japan
phone: 81-298-36-9189, fax: 81-298-36-9199
e-mail: abe@crc.riken.jp
PRIMERS
Sequencing : TJ
LIBRARY
Vector : pBACE3.6
R.Site 1 : EcoRI
R.Site 2 : EcoRI
Location/Qualifiers
1. 1542
/organism="Mus musculus molossinus"
/mol_type="genomic DNA"
/sub_species="moleosinus"
/db_xref="taxon:57486"
/clone="MSMG01-201G10.TJ"
/sex="male"
/tissue types="mixture of kidney and spleen"
/clone_lib="MSMG01 Mouse Male BAC Library"

Query Match 4.7%; Score 95.6; DB 9; Length 1542;
Best Local Similarity 44.3%; Pred. No. 2.2e-07;
Matches 420; Conservative 0; Mismatches 523; Indels 6; Gaps 1;

QY 750 ATAATAGTAGGCTACATGACATTAATTAATAGAGCTTAAATTAAGTCTACTATAA 809
DB 1103 ATAAGAAATATATAATAATAATAATAATAATAATAATAATAATAATAATA 1044
QY 810 ATAAGAGAGCTTAGTAAACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 869
DB 1043 AATAATAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA 984
QY 870 TTGTTGAGCTCATTTCTTTTAAAGTAATGTAAGTCTAAAGCAGACATAGAAATTA 929
DB 983 TATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA 924
QY 930 GTACAGGTTAAATCTTTTACAGAGTTTATTAATAACGAAATCAATTTTACATGCT 989
DB 923 ATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA 864
QY 990 CTCGGCTGCTATTATATAGGATCACTTACTGATCATCAATTA 1043
DB 863 TTAATTTATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 804
QY 1044 AAAACAAATTCATGAGATAATAATCTTCAATGAAGAGAGGACATCTCTTTTGA 1103
DB 803 AAAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 744
QY 1104 AAAACAAATAGTACTCCCTCGCTCCCTCTGAAATGTATACATATGGAATGGACGGAG 1163
DB 743 TAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA 684
QY 1164 ACTAAGAAAAATGTATAAGTAATGTAGTAGTAAAGAGAGAGAGAGAGAGAGGTTAA 1223
DB 683 AATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA 624
QY 1224 GTAGCGGAGCCCAATATATAATTTGATAGATTTAGAGAGAGTAGTTGAAAGTAGTGGT 1283

Db 623 ATAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 564
QY 1284 GGTGGGATTTTATATATAATAATAATAATAATAATAATAATAATAATAATAATA 1343
Db 563 TATATATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 504
QY 1344 TTGAGTGGGACATCCATAAAGAGAGTGTATAGATTAATTAATGGGACAGAGGAGTATA 1403
Db 503 TAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 444
QY 1404 CCTTATGATATATAATAATAATAATAATAATAATAATAATAATAATAATAATA 1463
Db 443 ATTATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 384
QY 1464 ATGATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 1523
Db 383 ATAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 324
QY 1524 TATAATTTTACAACTAGTATATATGAATAATCAATGATCTGAAAGCAATATATC 1583
Db 323 ATATATATATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 264
QY 1584 TTTGTAACACAGCGTTTCGGTCAATCGGAGTTTCATGTGATTCATAGTTTAAATATA 1643
Db 263 AATAATTTAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 204
QY 1644 AAGTAATTTTAAATTAATTTGTTTATTTTTCGTTTCAGAAATTTAAAAATA 1692
Db 203 AATAATATATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 155

RESULT 4
CNS0039G 1101 bp DNA linear GSS 03-JUN-1999
Drosophila melanogaster genome survey sequence TET3 end of BAC #
BACR08K10 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
AL063921
AL063921.1 GI:4941778
GSS.
Drosophila melanogaster (fruit fly)
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 1101)
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see <http://www.fruitfly.org> The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osogawa and
Aaron Mammose in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain y2; cn bw sp, the same strain used for the BDGP's
pi and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
Location/Qualifiers
1. 1101
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR08K10"

FEATURES
source

ORIGIN

| | | | |
|---------------------------|------|--|--|
| ORIGIN | | /clone lib="RPCI-98" /note="end : TET3" | |
| Query Match | | 4.6%; Score 95; DB 9; Length 1101; | |
| Best Local Similarity | | 20.2%; Pred. No. 2.8e-07; | |
| Matches 173; Conservative | | 325; Mismatches 348; Indels 9; Gaps 2; | |
| QY | 1200 | GAAAGAGAGAAAGTGGGTAAAGTAGCGGGACCCACCAATATATATAATTGATAGATTTA 1259 | |
| Db | 189 | GTWAKGGGGAATARRGGGGGGAARGGGGGGRGGGAARAGRARRGRRAAARAAA 248 | |
| QY | 1260 | GAAAAGTAGTGAAGAGTAGTGGGTGGGTGGGATTTTATATATATAAAAATTTACTATTTT 1319 | |
| Db | 249 | AARGRRARRGRRGRRGRRGRRGRRGRRGRRGRRGRRGRRGRRGRRGRRGRRGRRG 308 | |
| QY | 1320 | GAGAAAGTTTGGAAATGTATAGAAATGTAGTGGGACATCCATAAAAGAAAGTGTATAGAA 1379 | |
| Db | 309 | GARGRRRRRRRRRAAAGGRRRRRRRAGGRRRRRRRRRRRRRRRRRRRRRRRRRRRR 368 | |
| QY | 1380 | TTAAATGGGACAGGGAGTAATACCTTTATGATATATAAATTTTGTATTTTGATTTTC 1439 | |
| Db | 369 | RRGAGGGGRRGGGGGGGGGGGATATAWAWWWTTTTTTTTTTTAAWAAWAAATAA 428 | |
| QY | 1440 | ATAAGATTATAAATCTATGTATAATGATAATAATTTTAAATAATACTATATTAAT 1499 | |
| Db | 429 | TTWAAWAAWAAATAATTTWAAWAAWAAWATAWTTTATWAAWAAWAAWAAWAAW 488 | |
| QY | 1500 | TCTGATTAGTCGATACCGGCTTTTATAATTTTACAATCTAGTAGTAATATGAATAATCA 1559 | |
| Db | 489 | TTTWTTTTATWATAWTTTWWTTTAAWAAWAAWAAWAAWAAWAAWAAWATAATTTWT 548 | |
| QY | 1560 | GTTATCTGAAGCAATATATCTTTGTATAACAGCGTTCGCTCAATCGGAGTTCAT 1619 | |
| Db | 549 | WWTYTTTAAWATAAAMCMWAATYHYTYYTYYTYYTYYTYYTYYTYYTYYTYYTYY 608 | |
| QY | 1620 | GTGTATTCAATAGTTTAAATATAAAGTAAATTTAAATTAATGTGTTATTTTGTTCAG 1679 | |
| Db | 609 | HWYHTYAMWHMTWHTWAWWHWTYTAAYYYTYTCMY-----YHYMWHHAHAHA 662 | |
| QY | 1680 | AAATTTAAATAAATATTGAGCAGGGAAGTTTCAGGGCATCATTTGAGCAGCACTAGAC 1739 | |
| Db | 663 | AWTHTTWTWTHAYHWATYHYYYTYMYCAMCMCTHTCHCYYYHYHTAHHHTHHWYAHY 722 | |
| QY | 1740 | TGTTTGNACAAATGATGTCGGGTGATCATCTATGACCTTTCAACTCAAACTAGTGAATA 1799 | |
| Db | 723 | YMYWYMYWYMYCTACTYHYHHHHYHWYHTTWYAWAHMWWHHHAHYAAAAWAAWA 782 | |
| QY | 1800 | T--GCATTCTAGAATACATCTTTTCAAAATTTCAACAAACACACAGCTTTTAACTTTCTTC 1856 | |
| Db | 783 | TTHYHHHTTHYMHHTYMYHYMYTCCYMCYTHCWHYHTYATCTWTHHHMTWTHWYHH 842 | |
| QY | 1857 | AACGGATTGGAATCCTTTTCTAACTTTTAAATAAATAAATAAATGATTAATGTAATATT 1916 | |
| Db | 843 | TWHHHTTTHWAWHHTTWCWWNHAATTWATWACMCTWHHMHHMHHMHHMHHMHHMHH 902 | |
| QY | 1917 | TATCAACACCTCAACATTGATGTAGGTACTATAATAGGTCTCTTGCTGCTCTACTA 1976 | |
| Db | 903 | THMCMCHHHHCHCTCHHHHTYHMTCHWMMHMMHHMMHMMHMMHMMHMMHMMHMMH 962 | |
| QY | 1977 | TCATCATCAATCTTACACACAAACCTTGAGCTTAAATTTTCTACTTATCTTCAGCAA 2036 | |
| Db | 963 | CHYHMMHMYMCHYCTCTHTATTHYHMYMCTCYHYCTWHTTWTYAWWATAHAMTTAT 1022 | |
| QY | 2037 | TAACATCTTAATAT 2051 | |
| Db | 1023 | WWWMMHWAHWATWW 1037 | |
| RESULT 5 | | CL509408 | |
| LOCUS | | 1758 bp DNA linear GSS 01-APR-2004 | |
| DEFINITION | | SAIL_811_H11.v3 SAIL Collection Arabidopsis thaliana genomic clone | |

| | | | |
|--|-----|--|--|
| SAIL_811_H11.v3, genomic survey sequence. | | | |
| CL509408 | | | |
| CL509408.1 GI:46006728 | | | |
| GSS. | | | |
| Arabidopsis thaliana (thale cress) | | | |
| Arabidopsis thaliana | | | |
| Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids. | | | |
| 1 (bases 1 to 1758) | | | |
| Sessions,A., Burke,E., Presting,G., Aux,G., McElver,J., Patton,D., Dietrich,B., Ho,P., Bacwaden,J., Ko,C., Clarke,J.D., Cotton,D., Bullis,D., Snell,J., Miguel,T., Hutchison,D., Kimmery,B., Mittel,T., Katagiri,F., Glazebrook,J., Law,M. and Goff,S.A. | | | |
| A high-throughput Arabidopsis reverse genetics system | | | |
| Plant Cell 14 (12), 2985-2994 (2002) | | | |
| 22356987 | | | |
| 12468722 | | | |
| Contact: Sessions A | | | |
| Applied Trait Genetics | | | |
| Syngenta Biotechnology Inc. | | | |
| 3054 Cornwallis Rd., Research Triangle Park, NC 27709, USA | | | |
| Email: allen.sessions@syngenta.com | | | |
| ABRC Stock Number CS936276; T-DNA left border flanking sequences of Syngenta Arabidopsis Insertion Library (SAIL) lines are available through the Arabidopsis Biological Resource Center (ABRC). | | | |
| Sequences represent a pool of amplified genomic regions and not single contiguous sequences. | | | |
| Class: TDNA tagged. | | | |
| Location/Qualifiers | | | |
| 1..1758 | | | |
| /organism="Arabidopsis thaliana" | | | |
| /mol_type="genomic DNA" | | | |
| /ecotype="Columbia" | | | |
| /db_xref="taxon:3702" | | | |
| /clone="SAIL_811_H11.v3" | | | |
| /clone_lib="SAIL Collection" | | | |
| /notes="T-DNA left border sequences were isolated using a modified TAIL-PCR strategy" | | | |
| ORIGIN | | | |
| Query Match | | | |
| Best Local Similarity | | | |
| Matches | | | |
| 502; Conservative | | | |
| 0; Mismatches 957; Indels 7; Gaps 5; | | | |
| QY | 495 | AAATTGACAAACCCGACACACTCTATCCAGTACTACTTTTGGCCGATGCTTC 554 | |
| Db | 68 | ANANNAAAAAAAAAAAAAAAAANNNNNNNNNNNNNNNNNNNNNNNNNNNNATTTTTT 127 | |
| QY | 555 | TCAAAATGTTTTTATATGTAATAATGCCCATCCAAGGATAAGTAAATTCCTGTTTA 614 | |
| Db | 128 | NNAA 187 | |
| QY | 615 | ACCAGTTTGTATATATATGTTTACATTTACAAGAGGATATTCGTAATCTTTAGACG 674 | |
| Db | 188 | AANNNNNNNNNNAANN 247 | |
| QY | 675 | ACAAGAGCTTAGTCAAAAATGGACGCTGGTAAACAGCCTAGACTTGGTCACTGATAA 734 | |
| Db | 248 | ANAAAAANNAANNN 307 | |
| QY | 735 | TAGATAATTTGTAGTATAATATAGTAGGATCTCAATGACATTTAAATTTAGAGCTATTAA 794 | |
| Db | 308 | NNNAAAAAAAAAAAAAAAAAANNNNNAAAAAAAAAAAAAAAAAAAAAAAAATATATATT 367 | |
| QY | 795 | TTAAGTTACTATAATAAGAGAGGTTAGTAAACAGAAAGCAGGTAAACAAACAGAGCTTG 854 | |
| Db | 368 | TTTTATTANTANANAANN 427 | |
| QY | 855 | CTGCTGTGTTGTAGTTGTTGTGAGCTCATTTCTTTAAAAAGTAATGTAAACTGATCTAAA 914 | |
| Db | 428 | NNNAANNNNAAAAAAAAAAAAAAAAAANNNNNNNNNNNNNNNNNNNNNNNNNNN 487 | |

| | | | |
|----|------|--|------|
| Qy | 915 | GCACATAGAAAATTTAGTACAGGTTAAAACTTTTACAAGATTTTATATTAACAGGAATCA | 974 |
| Db | 488 | NNAAAAAANNNNNAANAAAAAANNNNAANNAANNAANNAANNAANNAANNT | 547 |
| Qy | 975 | TTTTTATAACATGTCTCTCGGCTGTCATTATAATAGGATCACTTACTGATCATCCATTAA | 1034 |
| Db | 548 | AANAAANAAAAAANNAANNAANNAANNAANNAANNNNNNAANNAANNAANNAANNN | 607 |
| Qy | 1035 | AACCTTGTAAAAAATAAATTCAGATAGATAAAATACTTACAATGAAAGAGACAAATGT | 1094 |
| Db | 608 | NNNNNNAAAAAANNNNNAANNAANNAANNAANNAANNAANNTAANNAANNNNAANNA | 667 |
| Qy | 1095 | CTCTTTGAAAAAACAATAGGTACTCCCTCGCTCCCTCTGAAATGTTATACATATGGATTG | 1154 |
| Db | 668 | ANANNNNAANNNNAANNNNNNAANNAANNAANNAANNAANNAANNAANNAANNA | 727 |
| Qy | 1155 | GACACGGAGACTAAGAAAAATGTTAAAGTAACTGTAGAGTAAAAAGAAAGAGAAAGAAA | 1214 |
| Db | 728 | AAAAAANANNNNNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAAN | 787 |
| Qy | 1215 | GTGGGTAAGGTAGCGGACCCACCAATATATAATTTGATAGATTTAGAAAAAGTAGTTGAAA | 1274 |
| Db | 788 | NAANNNNAANNNNNNNNAATAAANNNNAANNAANNAATAAATAAATAAATAATATAA | 847 |
| Qy | 1275 | GTACTGGGTGGTGGGATTTTATATATATAAAAAATTTACTATTTTGAGAAGTTTGGAAA | 1334 |
| Db | 848 | AATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA | 907 |
| Qy | 1335 | TGTATAGNAATTGAGTGGGACATCCATAAAGGAAAGCTGTATAGNAATTTAAATGGGACAG | 1394 |
| Db | 908 | TAITTTNNAANAATAATANAANAANTTATAAATAAATAAATAAATAAATAAATAAATA | 967 |
| Qy | 1395 | GGAGTAATACCTTTATGATATATAAAATTTTGTGTTATT-TTGATTTCATAAGATTATAAAT | 1453 |
| Db | 968 | TTTTTATATAAANNTTAATATATATATATATNTNAAATTTATNTNATTTAATAATAAT | 1027 |
| Qy | 1454 | CTATGT-TATAATGATAATAATAATTTTAAAAATAATACTATATTAATTT--CTGATTAGTC | 1510 |
| Db | 1028 | TTAANTANAANAANTAATAAANAANTAATAATAAATAATAATAATAATAATAATAATA | 1087 |
| Qy | 1511 | GATTACCGCCTTTTATAAATTTTACAATCTGAGTAAATATGAAATAATCAGTTATCTGAAA | 1570 |
| Db | 1088 | ATANATTAATAANNTNNNTTTTAANTATAAATAAANTTTTNTATATAATAAANTTTAATA | 1147 |
| Qy | 1571 | AGCAATAATAATCTTTGTGAAACAGCGTCCGTCAAATGGGAAGTTCATGTGATTCCAAT | 1630 |
| Db | 1148 | TAATAATAAANAATAAATANAANNTTAAANATTTAANATTTATAANNTTATAAAAAATTA | 1207 |
| Qy | 1631 | AGTTTAAATATAAAAGTA--AAATTTTAAATTTAAATTTGTTATTTTGTTCAGAAATTTAAA | 1688 |
| Db | 1208 | AATAATATNTAANTATTAATAAANTNTATATATAAATAATATATAANTTAANDATTTNT | 1267 |
| Qy | 1689 | ATAAATATTGAGCATGGGAAGTTACGGGCATCATTTGAGCAGCACATGACTGTTTGAAC | 1748 |
| Db | 1268 | ATTAANTATTAAATTTAAAAATAATAAANNTTAANTAAAAANTTANNNAAAAAATATATA | 1327 |
| Qy | 1749 | AATGTATGTCGGGTGATCATCTATGACCTTTCAACTCAAACTAGTGAATATGCAATCTTA | 1808 |
| Db | 1328 | TATATNNNTATNTTTATA-TTNTATAAATNNNTAATANNNAATATATATATATTTTNNNA | 1386 |
| Qy | 1809 | GAATACATCTTTTCAAATTTTCAACAAAAACACAGCTTTAACTTTTCTTTCAACGGATGGA | 1868 |
| Db | 1387 | AANTATNTNTATATNTATATTTTNTTATATTTATTTATATAAATATATATTTAATAAAAA | 1446 |
| Qy | 1869 | TCCTTTTCTAAACTTTTTTAAAAATAAAAAAATGCAATTTATGTAATATTTTATCAACACCT | 1928 |
| Db | 1447 | ANTAATAAATAAANAATAATAAAAAAATAAATAAATAAATAAATAAATAAATAAATAA | 1506 |
| Qy | 1929 | AACATTGATGTTAGGTACTATAAAT | 1954 |
| Db | 1507 | AATAAANAATATNTTTTATNTNANAT | 1532 |

```

RESULT 6
CG753083/c
LOCUS      P048-1-C01.2a Ppa EcoRI BAC Library Pristionchus pacificus genomic, GSS 24-OCT-2003
DEFINITION      Pristionchus pacificus genomic sequence.
ACCESSION      CG753083
VERSION        CG753083.1 GI:37977199
KEYWORDS       GSS.
SOURCE         Pristionchus pacificus
ORGANISM       Pristionchus pacificus
REFERENCE      1 (bases 1 to 1896)
AUTHORS        Srinivasan,J., Sinz,W., Jesse,T., Wiggers-Perebolte,L., Jansen,K.,
                Buntjer,J., van der Meulen,M. and Sommer,R.J.
TITLE          An integrated physical and genetic map of the nematode Pristionchus
                pacificus
JOURNAL        Mol. Genet. Genomics 269 (5), 715-722 (2003)
MEDLINE        22835951
PUBMED         12884007
COMMENT        Contact: Sommer RJ
                Evolutionary Biology
                Max-Planck-Institute for Developmental Biology
                Spemannstr. 37-39, Tuebingen D-72076, Germany
                Tel: 00497071601371
                Fax: 00497071601498
                Email: ralf.sommer@uebingen.mpg.de
                Class: BAC ends.
FEATURES             Location/Qualifiers
     source           1..1896
                     /organism="Pristionchus pacificus"
                     /mol_type="genomic DNA"
                     /strain="California"
                     /db_xref="taxon:54126"
                     /clone_lib="Ppa EcoRI BAC Library"
                     /note="The library was generated by a partial digest of
                           the genomic DNA with EcoRI and cloning into the BAC
                           vector."
ORIGIN
Query Match      4.6%; Score 93.6; DB 9; Length 1896;
Best Local Similarity 44.1%; Pred. No. 5.1e-07;
Matches 520; Conservative 0; Mismatches 645; Indels 15; Gaps 4;

Qy 531 ACTATACCTTTGGCCGAATGCTTCTCAAAATGTTTTTATATGTATGTAATAATATGCCCATCC 590
Db 1933 AATTATATATAAAAAAATNTATATAAAAAATAAATAATATATATATAAAAAAATTTAAA 1774

Qy 591 AAGGATAAGTAAATATCCCGTTTAAACAGATTTGTTTAATATATATATGTTTACACTTCAAGA 650
Db 1773 AATAAAAAATTTAAATATATAAAATAATAAATAAATAAATAAATAAATAAATAAATAAATA 1714

Qy 651 GGATATTCGTAATACCTTTTAGACGACGACGACTTAGTCAAAAATGGACGCTGGTAAC 710
Db 1713 TATTAAATATTNATATAAAATATAAAAAATNTAAATTTATATATANAATAATAATTTAAAAATNA 1654

Qy 711 AGCCTAGACTTGGTCACTGATAATAAGATAATTTGTAGTATAATATAGTAGGATCTACAA 770
Db 1653 ATAAATATTAATTANNAAAAAATAATATTTTNNAAATTTTATATATATATAATAATATAT 1594

Qy 771 TGCATTTAAATTTAGAGCTATTAATTAAGTTACTAATAAATAAGAGAGGTTAGTAAACAG 830
Db 1593 TTAATAATATNATATAAAATAATTATTATATATTTTATATATTTTATTTTAAATATATAT 1534

Qy 831 AAGCAGGTAAAAACAAGAGCTTGCTGCTGTGTGTTTAGTTGTTGTGAGCTCATTTCTTT 890
Db 1533 AATAATTTTAAATATTTTATTTTAAAAAATAATATATANATAAAAAATATATTTTATTTTATTT 1474

Qy 891 AAAAGTAATGTAACTGATCTAAAG -CACATAGAAATTTAGTACAGGTAAAACTTTTAC 949
Db 1473 TTTTTTTATNTAAATATTTATTTATTTATAAATAAATTTNAAATTAATAAAAAAATATATAN 1414

Qy 950 AAGAATTTATTTAAACGAAATCATTTTTTATAACATGCTCTCGGCTGTCAATTAATAG 1009

```

```
Db      1413 TAAATTTTATTTTAAATATTTAAATATATTTAAATTTNANNAATATAAATATATAT 1354
Qy      1010 GGATCACTTACTGATCATCCATTAACCTTGTGTTAAACAAATTCATGAGATATAAATAT 1069
Db      1353 AAAAAATATNATAATTAATAATTAATAATATATTTTAAAAATTTTAAATTAATAATAT 1294
Qy      1070 CTTACAAATGAAAGAGGCAATGCTCTTTGAAAAACAACATAGGTACTCCCTCGTCC 1129
Db      1293 AAAAAATAAAAAATAAATAATTTATATAATAATAATAATAATAATAATAATAATAA 1234
Qy      1130 CTCTGAAATGTATACATATGGAATGGACACGGGAGACTAAGAAAAATGTTATAAGTAATGT 1189
Db      1233 AAAAAATATATATAATTAATATATATATAATAAAAAATAAATAATTTAAATAAATTT 1174
Qy      1190 AGAGTAAAAAGAAAGAAAGAAAGAGTGGGTAAAGTAGGGGACCCACCAATATATAAT 1249
Db      1173 ATAAATATATATATAATAATAAAAAATAAATATAA-----AAAAAATAAATAT 1123
Qy      1250 GATAGATTTAGAAAGTAGTTGAAAGTAGTGGGTGGGATTTTATATATATAAATAAT 1309
Db      1122 TATAAATTTAAAAATTAATAATNTATTTTATTAATTTAAATATATAATAATTTATTTAT 1063
Qy      1310 TTACTATTTTGAGAAAGTTTGAATGCTATAGAAATTCAGTGGGACATCCATAAAGGAAA 1369
Db      1062 TTATATAATATTAATATATATTTTATATATATATATAATAAATAAATTTATAATAAT 1003
Qy      1370 GTGTATAGAAATTAATGGGACAGAGGAGTAATACCTTTATGATATATAAATTTTGTGA 1429
Db      1002 TTTTAAATAAATAAATAANATATTTTATTTATTAATTTTATTAATAATATATAAATAA 943
Qy      1430 TTTTGTATTTTCATAGATTAATAATCTA--TGTTATATGATATATAAATTTTAAATAA 1487
Db      942 AATTTTATTTTAAATAAAAAANAATTAATTTATTAATAATATTTATAATAATAAATAA 883
Qy      1488 TACTATATT---AATCTGATTTAGTCGATTCAGCTACCGCTTTTATATTTTACAATCTGAGT 1544
Db      882 AAATAAATTAATAAAAAAATAAATAAATAAATAAATAAATAATTTATATATAATTT 823
Qy      1545 AATATGAATAAATCAGTTATCTGAAAAGCAATAATATCTTTGTAAAAACAGCGTTCGGTC 1604
Db      822 ATNATNAAAAAATTAATAAATTAATAATATAAATAATTTTAAATTTTAAATATTTAA 763
Qy      1605 AATGGGAAGTTCATGTGATTTCAATAGTTTAAATAAGTAATAATTTTAAATTAATTCG 1664
Db      762 ATTATATAATAATAAATAAATAAATTTATAATAATAAATAAATAAATAAATAAATAA 703
Qy      1665 TTATTTTGTGTCAGAAATTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1704
Db      702 ATATAATAATAATAATTTTATTTTATTTATATATTTATTAAT 663
```

```
RESULT 7
CC262481/c
LOCUS      1202 bp      DNA      linear      GSS 13-MAY-2003
DEFINITION CH261-167M9_Sp6.1 CH261 Gallus gallus genomic clone CH261-167M9,
            genomic survey sequence.
ACCESSION  CC262481
VERSION    CC262481.1 GI:30607397
KEYWORDS   GSS.
SOURCE     Gallus gallus (chicken)
ORGANISM   Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE  1 (bases 1 to 1202)
AUTHORS   Kremitzki, C., Higginbotham, J., Wylie, K., Carter, J., McPherson, J.,
            Warren, W., Graves, T., Mardis, E. and Wilson, R.
TITLE     Gallus gallus BAC End Reads
JOURNAL   Unpublished (2003)
COMMENT   Contact: Richard K. Wilson
            Genome Sequencing Center
            Washington University School of Medicine
```

Email: submissions@wustl.wustl.edu
Insert Length: 182000 Std Error: 0.00
Seq primer: Sp6 ATTTAGGTGACACTATAG
Class: BAC ends
High quality sequence start: 30
High quality sequence stop: 105.
Location/Qualifiers
1. .1202
/organism="Gallus gallus"
/mol_type="genomic DNA"
/strains="Red Jungle Fowl"
/db_xref="taxon:9031"
/clone="CH261-167M9"
/sex="female"
/cell_line="UCD001, inbred 256"
/clone_lib="CH261"
/note="Vector: pTARBAC2.1; Site 1: EcoRI; Site 2: EcoRI;
CH261 Female Chicken library - for library and clone
ordering information: http://www.chori.org/bacpac"

ORIGIN

Query Match 4.5%; Score 93; DB 8; Length 1202;
Best Local Similarity 47.3%; Pred. No. 6.6e-07;
Matches 393; Conservative 0; Mismatches 420; Indels 17; Gaps 4;
Qy 886 TCTTTTAAAGTAAATGTAAGCTGATCTAAAGCACATAGAAAATTTAGTACAGGTTAAAACTT 945
Db 1174 TATNANAATAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 1115
Qy 946 TTACAAGAAATTTATATTAATAACGAAATCATTTTAAACATGTCTCTCGCGCTGTCATTATA 1005
Db 1114 ATAAAAATTTTATATATAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1055
Qy 1006 ATAGGATCATTCTGATCATCTTAAACCTTGTGTTAAACAAATTCATGAGATATAA 1065
Db 1054 AAAATAATATAATAATAATAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1003
Qy 1066 ATATCTTCAATGAAAAGAGGACCAATGCTCTTTGAAAAAACAATAAGTACTCCCTCC 1125
Db 1002 ATATNTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 943
Qy 1126 GTCCCTCTGAAATGTATACATATGGATTGGACACGGAGACTAAGAAAAATGATATAAGTA 1185
Db 942 ATATTTTAAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA 883
Qy 1186 ATGTAGATGAAAAGAGAGAGAGAGAGAGTGGGTAAAGTAGCGGACCCACCAATATAT 1245
Db 882 ATTAATAATTTATAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA 823
Qy 1246 AATTGATAGATTTAGAAAAGTAGTTGAAAAGTAGTGGGTGGGATTTTATATATATATAA 1305
Db 822 AAAAAATAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA 763
Qy 1306 AAATTTACTATTTTGAGAAAAGTTTGAATGATATAGAATTTGAGTGGGACATCCATAAAG 1365
Db 762 TAAATAATTTAAATATAATAATAATAATAATAATAATAATAATAATAATAATAATAAAT 703
Qy 1366 GAAAGTGATAGAAATTAATGGACAGAGGAGTAATACCTTTATGATATATAAATTTT 1425
Db 702 AAAAA--TAAAAATNTAAATTTAAATAATAATAATAATAATAATAATAATAATAAATTT 645
Qy 1426 GTATTTTGTATTTTCAAGATTTATAAATCTATGTTATATAATGATATAATAATTTTAA 1485
Db 644 TATTTATATATAATAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA 585
Qy 1486 AATACATATATTAATCTGATTAGTCGATTACCGCCCTTTTATATTTTACAATACTCAGTA 1545
Db 584 AAAAAAATTTTATTTTATATATAATAATAATAATAATAATAATAATAATAATAATAA 531
Qy 1546 ATATGAATAAATCAGTTATCTGAAAAGCAATAATAATCTTTGTAAACAGCGTTCGGTCA 1605
Db 530 ATATATAATAATTTTAAAAATAATTAATAATAATAATAATAATAATAATAATAATAATA 471

Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

```
FEATURES
  source
    1..1101
      /organism="Drosophila melanogaster"
      /mol_type="genomic DNA"
      /db_xref="taxon:7227"
      /clone="BACR29B23"
      /clone_lib="RPCI-98"
      /notes="end : T7"

ORIGIN
  Query Match      4.4%; Score 90.8; DB 9; Length 1101;
  Best Local Similarity 34.7%; Pred. No. 1.7e-06;
  Matches 197; Conservative 106; Mismatches 264; Indels 1; Gaps 1;

Qy 1130 CTCGAAATGTATACATATGATGGATGGACACGGACATGAAGAAAAATGTATAAAGTAATGT 1189
    ||| | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 454 YTCCTAHTTMMWMMWAAATWTWAAAWAAATTTATWAAATWAAAWAAWMMWATTTT 513

Qy 1190 AGAGTAAAGAAAGAGAAAGAGGAGTGGGTAAAGTAGCGGGACCCACCAATATAAT 1249
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 514 WMMWTATTTTMMWMTWTWATAAATAAATAAATAAATAATTTAAAWAAATATAAAT 573

Qy 1250 GATAGATTTTGAAGAAGTAGTTGAAGTAGTGGGTGGGTGATTTTATATATAAATAAT 1309
    ||| | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 574 TAWAAWTATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTT 633

Qy 1310 TTACTATTTTGAGAAAGTTTGAATGTATAGAAATGAGTGGGACATCCATTAAGAGAA 1369
    ||| | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 634 TTAATAATTTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 693

Qy 1370 GTGTATAGATTAATAGGGACAGGGAGTAATACCTTTTATGATATATAATTTTGTGA 1429
    ||| | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 694 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 753

Qy 1430 TTTTGATTTTCAAGATATAATATCTATGTATAATGATAATATAATTTTAAAAATAA 1489
    ||| | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 754 ATATWATATATATTTTMMWMTWTWAAWMTATATWATWATWATWAAWAAWATAATA 813

Qy 1490 CTATATTAATCTGATTAAGTACCGCTTTTATAATTTTACAACTAGTAATAT 1549
    ||| | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 814 -TAAWAAAWAAWATAWATWATATWAAATAWAAWAAATAATWTAATWAAATWATAA 872

Qy 1550 GAATAAATCAGTTATCTGAAGAAGCAATAATATCTTTGTAAACACAGGTCGCTCAATG 1609
    ||| | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 873 WAAAAATAWAAWTTTMMWMTWTWAAWATATAWAAWATAWAAWAAAAAATAAATA 932

Qy 1610 GGAAGTTCATGTATCAATAGTTTAAATAAAGTAATTTTAAATTAATTTTGTATT 1669
    ||| | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 933 ANAWMTWATATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 992

Qy 1670 TTTGTTTCAGAAATTTAAAAATAAATTAAT 1697
    ||| | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 993 ATWWTATWATATATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT
```

```
RESULT 10
CL118721
LOCUS
DEFINITION
  ISB1-72J8 T7.1 ISB1 Xenopus tropicalis genomic clone ISB1-72J8,
  genomic survey sequence.
ACCESSION
  CL118721
VERSION
  CL118721.1 GI:40612356
KEYWORDS
  GSS.
SOURCE
  Xenopus tropicalis (western clawed frog)
  ORGANISM
    Xenopus tropicalis
```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae; Xenopodinae; Xenopus; Silurana.

REFERENCE
 1 (bases 1 to 1608)
 Krenitzki,C., Carter,J., McPherson,J., Warren,W., Graves,T.,
 Mardis,E. and Wilson,R.
 A physical map of the xenopus tropicalis genome
 Unpublished (2003)

COMMENT
 Contact: Richard K Wilson
 Genome Sequencing Center
 Washington University School of Medicine
 Email: submissions@watson.wustl.edu
 Insert Length: 75000 Std Error: 0.00
 Seq primer: T7 TAATACGACTCACTATAGG
 Class: BAC ends
 High quality sequence start: 252
 High quality sequence stop: 345.

FEATURES
 source
 1..1608
 /organism="Xenopus tropicalis"
 /mol_type="genomic DNA"
 /db_xref="taxon:8364"
 /clone="ISB1-72J8"
 /clone_lib="ISB1"
 /note="Vector: pBelobAC11; ISB-1 Xenopus tropicalis BAC
 Library Segment 1"

ORIGIN
 Query Match 4.4%; Score 90.4; DB 9; Length 1608;
 Best Local Similarity 45.6%; Pred. No. 2e-06;
 Matches 446; Conservative 0; Mismatches 516; Indels 16; Gaps 3;

```
Qy 730 ATAAATAGATAATTTCTAGTATAATATAGTAGTCTACAAATGACATTTAAATTTAGAGCT 789
    ||| | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 338 AAAAAAATAATTTTAAAAATAAATAATAATAATTTAAAAATAAATAAATAAATAATAA 397

Qy 790 ATTAATTAAGTTACTAATAATAAGAGAGGTTAGTAAACAGAAAGCAGGTAAAAACAAGA 849
    ||| | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 398 ATAAATTAATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 457

Qy 850 GCTTGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 905
    ||| | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 458 AAAAAAATAATTTTAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 517

Qy 906 ----TGATCTAAGCACATAGAAATTTAGTACAGTTTAAACCTTTTACAGAAATTTATAT 961
    ||| | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 518 TTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 577

Qy 962 TAAACGAAATCATTTTATAACATGTCTCTCGGCTGTCTATTAATAAGGATCAGTTACT 1021
    ||| | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 578 AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 637

Qy 1022 GATCATCCATTAACCTTTGTTTAAACAAATTCATGAGATATAATATCTTACAAATGAA 1081
    ||| | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 638 TAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 695

Qy 1082 AGAAGACAAATGTCCTTTGAAAAACAATAAGTAGTACTCCCTCCGCTCGCAATGTA 1141
    ||| | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 696 ATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 755

Qy 1142 TACATATGATTTGGACACGAGACTAAGAAAAATGTATAAGTAATGTAGAGTAAAAAGA 1201
    ||| | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 756 TATATAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 815

Qy 1202 AAGAGAAAGAAAGTGGGTAAAGTAGCGGAGCCCAATATATAATATAGATTTAGA 1261
    ||| | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 816 ATAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 875

Qy 1262 AAGTAGTTGAAGTAGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGG 1321
    ||| | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 876 AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 935

Qy 1322 GAAAGTTTGAATGTATAGAAATTTAGTGGGACATCCATAAAGGAAAGTGTATAGAAAT 1381
```


1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 (E-mail: hattori@gsc.riken.jp, URL: http://hgp.gsc.riken.go.jp/,
 Tel: 81-45-503-9111, Fax: 81-45-503-9170)
 Clones are derived from the mouse BAC library MSMg01. For BAC
 library availability, please contact Kuniyu Abe (abe@rtc.riken.jp).
 Teakuba Institute, Bio Resource Center,
 The Institute of Physical and Chemical Research (RIKEN) 3-1-1
 Koyadai, Teakuba, 305-0074 Japan
 phone: 81-298-36-9189, fax: 81-298-36-9199
 e-mail: abe@rtc.riken.jp

PRIMERS

Sequencing : T7

LIBRARY

Vector : pBACe3.6

R.Site 1 : EcoRI.

R.Site 2 : EcoRI.

Location/Qualifiers

1. .1268

/organism="Mus musculus molossinus"

/mol_type="genomic DNA"

/sub_species="molossinus"

/db_xref="taxon:57486"

/clone="MSMg01-142102.T7"

/sex="male"

/tissue_type="mixture of kidney and spleen"

/clone_lib="MSMg01 Mouse Male BAC Library"

ORIGIN

Query Match 4.3%; Score 88.8; DB 9; Length 1268;
 Best Local Similarity 45.5%; Pred. No. 3.9e-06;
 Matches 471; Conservative 0; Mismatches 551; Indels 14; Gaps 4;

QY 1030 ATTAAACCTTGTAAACAAATCAATGAGATAAATATCTTACAAATGAAAGAGGAC 1089
 DB 1215 ATTAATAATNATAAATAAATAATATATAAATAAATAAATAAATAAATAAAT 1156
 QY 1090 AATGCTCTTTGAAAACAAATAGGTACTCCCTCGCTCGCTGAAAT---GTATACAT 1146
 DB 1155 ANTAATATATATAAATTTAAATAATAAATAAATAAATAAATAAATAAATAA 1096
 QY 1147 ATGGATTGGACACGGAGACTAAGAAAATGTATAAAGTATGTAGAGTAAAGAAAGAG 1206
 DB 1095 ATTAATATATATTAATAATAATATTAATAATAAATAAATAAATAAATAAAT 1036
 QY 1207 AAGAAAAGTGGTAAAGTAGCGGACCCCAATATATAAATTAATGATGATTAAGAAAGT 1266
 DB 1035 AATAAATAATAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAAT 976
 QY 1267 AGTTGAAGTAGTGGGTGGGATTTTATATATAAATAAATAAATAAATAAATAA 1326
 DB 975 ATTTATAAATAATAATATTTATAAATAAATAAATAAATAAATAAATAAATAA 916
 QY 1327 TTTTGAATGTATAGATTGAGTGGGACATCCATAAAGGAAAGTGTATAGAAATTAATG 1386
 DB 915 ATAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 856
 QY 1387 GGCAGAGGAGTAATACCTTTATGATATATAAATAAATAAATAAATAAATAAATAA 1446
 DB 855 TTAATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 796
 QY 1447 TATAATCTATGTTATATATATAAATAAATAAATAAATAAATAAATAAATAAATA 1506
 DB 795 ATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 736
 QY 1507 AGTCGATTACCGCTTTTATAATTTTACAATCTGAGTATATGAATAAATCAGTTATCT 1566
 DB 735 AATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 676
 QY 1567 GAAAAGCAAT---AATATCTTTGTAAACACGCGTTCGGTCAATGGGAAGTTTCATGTGT 1623
 DB 675 AATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 616
 QY 1624 ATCAATAGTTTATAAATAAAGTAATTTTAAATTAATTTGTTATTTGTTTCAGAAAT 1683

DB 615 ATATTATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 556
 QY 1684 TTAATAATAAATATTGAGCATCGGAGTTCACGGGCATCATTTGA---GCAGCCTAGACT 1740
 DB 555 ATAAATAATATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 496
 QY 1741 GTTTCGAACAATGATGTCGGGTGACATCTATGACCTTTCAACTCAACTAGTGAATAAT 1800
 DB 495 ATATTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 436
 QY 1801 GCATTCTAGATACATCTTTTCAAAATTTCAACAAACACACAGCTTTTAACTTTTCTT 1855
 DB 435 AGTAAATAAATAAATAGGTATTTAATAAATAAATAAATAAATAAATAAATAA 376
 QY 1856 CAACGATTGGAATCCTTTTCTAAATCTTTTAAATAAATAAATAAATAAATAA 1915
 DB 375 AAAAAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 316
 QY 1916 TTATCAACACCTCAACATTTGATGCTAGCTACTATAAATAAGTCTCTTGGTCTACT 1975
 DB 315 TAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 256
 QY 1976 ATCATCACATCAATCTTACACCAACACCTTGCAGCTTAATTTTCTTACTTATTTCTCAGCA 2035
 DB 255 TAATTATATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 196
 QY 2036 ATAACATTTCTAATAAT 2051
 DB 195 ATAAATAAATAAATAAAT 180

RESULT 13
 CG757503/c
 LOCUS
 DEFINITION
 CG757503 1392 bp DNA linear GSS 24-OCT-2003
 P052-4-C08.za Ppa EcoRI BAC Library Pristionchus pacificus genomic,
 genomic survey sequence.
 CG757503
 CG757503.1 GI:37986131
 GSS.
 Pristionchus pacificus
 Pristionchus pacificus
 Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
 Neodiplogasteridae; Pristionchus.
 1 (bases 1 to 1392)
 Srinivasan,J., Sinz,W., Jesse,T., Wiggers-Perebolte,L., Jansen,K.,
 Buntjer,J., van der Meulen,M. and Sommer,R.J.
 An integrated physical and genetic map of the nematode Pristionchus
 pacificus
 Mol. Genet. Genomics 269 (5), 715-722 (2003)
 22835951
 12884007
 Contact: Sommer RJ
 Evolutionary Biology
 Max-Planck-Institute for Developmental Biology
 Spemannstr. 37-39, Tuebingen D-72076, Germany
 Tel: 00497071601371
 Fax: 00497071601498
 Email: ralf.sommer@tuebingen.mpg.de
 Class: BAC ends.

FEATURES
 source
 1. .1392
 /organism="Pristionchus pacificus"
 /mol_type="genomic DNA"
 /strain="California"
 /db_xref="taxon:54126"
 /clone_lib="Ppa EcoRI BAC Library"
 /note="The library was generated by a partial digest of
 the genomic DNA with EcoRI and cloning into the BAC
 vector."

ORIGIN
 Query Match 4.3%; Score 88.8; DB 9; Length 1392;

Best Local Similarity 43.8%; Pred. No. 3.9e-06;
Matches 543; Conservative 0; Mismatches 672; Indels 24; Gaps 3;

Qy 481 ATGCTCAGCCATCAAAATTTGACAAACCCGACACACACTCTCCACGCTACTATCTTT 540
Db 1359 ATATAAAATATATAAATATATAAATAAATAAATAAATAAATAAATAAATAAATAA 1300
Qy 541 TGGCGGAATGCTCTCAAAATGTTTTTATATATATATATATATATATATATATATAT 600
Db 1299 TAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1240
Qy 601 AAAATTCGCGTTTAAACAGTTTGTATATATATATATATATATATATATATATATAT 660
Db 1239 ATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAAT 1180
Qy 661 AATACTTTTACGACGACGAGACTTAGCTCAAAATGACGCTGCTGTAACAGCCTAGACT 720
Db 1179 AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1120
Qy 721 TGGTCACTGATAAATAGATAAATGTTAGTATATATATATATATATATATATATATAT 780
Db 1119 AAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1060
Qy 781 ATTAGAGCTATTAATTAACTTAAATAAATAAAGAGAGGTTAGTAAACAGAAACAGGTA 840
Db 1059 AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1000
Qy 841 AAAACAAGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
Db 999 AATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAAT 940
Qy 901 TAACTGATCAAGACATAGA-----AATTTAGTACAGGTTAAAACTTTTAC 949
Db 939 AATATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 880
Qy 950 AAGAAATTTATTAACGAAATCATTTTATAACATGCTCTCGGCTGCTCATTAATAATAG 1009
Db 879 AATAAATTTATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 820
Qy 1010 GGATCACTTACTGATCATCAATTAACCTTGTGTTAAACCAATTCATAGATAAATAAT 1069
Db 819 ATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 760
Qy 1070 CTTCAATGAAGAGGACAAATGCTCTTTGAAAAACAATAAGTAGTACTCCCTCCGCTC 1129
Db 759 ATTATTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 700
Qy 1130 CTCTGAATGTATACATATGGATTGGACGAGACTAAGAAATAATGATAAAGTATATGT 1189
Db 699 AAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 640
Qy 1190 AGAGTAAAAAG 1249
Db 639 AAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 580
Qy 1250 GATGATTTGAAAGAGTGTGAAA-----GTAGTGGGTGGGTGGGATTTTATATATAA 1305
Db 579 AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 520
Qy 1306 AAATTTACTATTTTGAGAAAGTTTGAATGTATAGAAATTTGAGTGGGACATCCATAAAG 1365
Db 519 ATTATATATATTTATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 460
Qy 1366 GAAAGTGTATAGATTAATGAGGACAGAGGAGTAATACCTTTTATGATATATAAATTTT 1425
Db 459 TAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 400
Qy 1426 GTTATTTTGTATTCATAGATTTAATAATCTATGTTAATAATGATAAATAAATTTTAA 1485
Db 399 AATAATTTATTTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 340
Qy 1486 AATACATATAATTTCTGATTAGTCGATTAACCGCTTTTATAATTTTACATACTAGTA 1545

Db 339 AAATATATATAATTAATTTATATATTTATTTATTTATTTATTTATTTATTTATTTAT 280
Qy 1546 ATATGATATAAATCAGTTTATCTGAAAACGAAATATATCTTTGTAAAAACAGCGTTTCGGTCA 1605
Db 279 TAAAAATTTTATATAAATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 220
Qy 1506 AATGGGAAGTTCATGCTGATTTCAATAGTATTTTATATATAAAGTAAATTTTAAATTAATGT 1665
Db 219 AAATATATATATAATTTT-----TAATAAATAAATAAATAAATAAATAAATAAATAA 169
Qy 1666 TATTTTGTCTTCAAAATTTTAAATAAATAAATTTATTGAGCAT 1704
Db 168 TATTTATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAAT 130

RESULT 14
CC253231 1277 bp DNA linear GSS 13-MAY-2003
LOCUS CH261-180N11_RM1.1 CH261 Gallus gallus genomic clone CH261-180N11,
DEFINITION genomic survey sequence.
CC253231
VERSION CC253231.1 GI:30589981
KEYWORDS GSS.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 1277)
Kremitzki, C., Higginbotham, J., Wylie, K., Carter, J., McPherson, J.,
Warren, W., Graves, T., Mardis, E. and Wilson, R.
Gallus gallus BAC End Reads.
Unpublished (2003)
Contact: Richard K. Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submission@wustl.edu
Insert Length: 182000 Std Error: 0.00
Seq primer: RM1 TAGCACTCACTATAGGAGA
Class: BAC ends
High quality sequence start: 24
High quality sequence stop: 85.

FEATURES
Location/Qualifiers
1..1277
/organism="Gallus gallus"
/mol_type="genomic DNA"
/strain="Red Jungle Fowl"
/db_xref="taxon:9031"
/clone="CH261-180N11"
/sex="female"
/cell_line="UCD001, inbred 256"
/clone_lib="CH261"
/note="Vector: pTARBAC2.1; Site 1: EcoRI; Site 2: EcoRI;
CH261 Female Chicken library - for library and clone
ordering information: http://www.chori.org/bacpac"

ORIGIN
Query Match 4.3%; Score 88; DB 8; Length 1277;
Best Local Similarity 43.8%; Pred. No. 5.4e-06;
Matches 475; Conservative 0; Mismatches 600; Indels 6; Gaps 3;

Qy 619 GTTTGTTTAAATATATATGTTTACACTTACAGAGAGATATTCGTAAATACTTTTAGACGCAA 678
Db 41 GT 100
Qy 679 GAGACTTAGGTCAAAAATGAGCGCTGGTAAACAGCCCTAGACTTGGTCACTGATATAATAGA 738
Db 101 GAACCTTTAAAAAANNCNCAAAAAAATACATATTTTATAAAAAATAATATTAAAAATAATA 160
Qy 739 TAATTTGTTAGTATAATATAGTAGAATCTACAATGACATTAAATTTAGAGCTATTAAATTA 798
Db 161 TATTTTAAACAATTTTGTGAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAT 220

| | | | | | | | | | | | | | | | | |
|----|------|-----|---------|-------|--------|--------|--------|-------|-------|------|------|--------|------|------|-------|-----|
| Qy | 799 | GT | TACTAAT | AAAA | TAAG | GAGGGT | TAGTAA | CAGAA | CGAGG | TAA | AAA | CAAGAG | CTT | CGT | CG | 858 |
| | | | | | | | | | | | | | | | | |
| Db | 221 | AT | TAAAT | TAAAT | TATAAT | TAAAT | TAAAT | TAT | TAT | TAT | TAT | TAT | TAT | TAT | 280 | |
| | | | | | | | | | | | | | | | | |
| Qy | 859 | TG | TGTT | TTAG | TTT | TG | TGAG | CT | CA | TTT | CTTT | TAAAG | TAA | TG | TAAAG | 918 |
| | | | | | | | | | | | | | | | | |
| Db | 281 | AT | AAAT | TAT | TAAT | TAAAT | TAAT | AT | AT | AT | AT | TAAT | TAAT | TAAT | 340 | |
| | | | | | | | | | | | | | | | | |
| Qy | 919 | AT | GAAA | AT | TTAG | TACAG | GGT | TTAAA | ACT | TTTT | ACA | GAA | TTT | TAT | 978 | |
| | | | | | | | | | | | | | | | | |
| Db | 341 | TAT | AAAT | TTA | TAA | TAT | TAT | TAA | TAT | TAA | TAT | TAA | TAT | TAA | 400 | |
| | | | | | | | | | | | | | | | | |
| Qy | 979 | ATA | CAT | GT | CT | CT | CGG | TG | TCAT | TAT | TAAT | AGG | AT | CAT | 1038 | |
| | | | | | | | | | | | | | | | | |
| Db | 401 | ATT | TAT | TAAT | AAAA | TAT | AAAT | TAT | AA | TAA | TAA | TAA | TAA | TAA | 460 | |
| | | | | | | | | | | | | | | | | |
| Qy | 1039 | TT | GT | TAAA | CA | AA | TT | CA | AT | CT | T | CA | AT | CA | 1098 | |
| | | | | | | | | | | | | | | | | |
| Db | 461 | TTT | TAT | TAT | TTT | TAT | TAA | TTT | TAT | AAAA | TT | T | AA | TT | 520 | |
| | | | | | | | | | | | | | | | | |
| Qy | 1099 | TT | GAAAA | AAAA | CA | AA | TAG | TAC | T | CC | T | CG | T | CC | 1158 | |
| | | | | | | | | | | | | | | | | |
| Db | 521 | AT | AT | TAA | CT | TAT | AT | TTT | TAA | TAA | TAT | TAT | TAA | TAT | 580 | |
| | | | | | | | | | | | | | | | | |
| Qy | 1159 | CG | GA | - | GACT | AG | AAAA | AT | GT | AT | AA | AG | TG | TAG | 1217 | |
| | | | | | | | | | | | | | | | | |
| Db | 581 | TAA | AT | TTT | TAA | AT | AAAA | TT | TAA | AT | TTT | TAA | AT | TAT | 640 | |
| | | | | | | | | | | | | | | | | |
| Qy | 1218 | GG | TAA | AG | T | AG | GG | AC | CC | CA | TAT | AT | T | AT | 1277 | |
| | | | | | | | | | | | | | | | | |
| Db | 641 | AT | AT | AA | TAT | AAAA | TT | TAA | AA | TAT | TA | AAAA | AA | TAT | 700 | |
| | | | | | | | | | | | | | | | | |
| Qy | 1278 | GT | GG | T | GG | T | GG | AT | TTT | TAT | TAT | AAAA | TT | T | 1337 | |
| | | | | | | | | | | | | | | | | |
| Db | 701 | AT | TAA | TTT | TAA | TTT | TAT | AT | AT | TAA | TAT | TAA | TAT | TAA | 760 | |
| | | | | | | | | | | | | | | | | |
| Qy | 1338 | AT | GAA | TT | G | AG | T | GG | GA | CA | T | CC | AT | AA | 1397 | |
| | | | | | | | | | | | | | | | | |
| Db | 761 | AA | TAT | AT | AT | AT | TTT | TAA | TAA | | | | | | | |

| | | | | | | |
|------------|------------|-------------------------|------------------------------|-----|--------|-----------------|
| RESULT 15 | AG346840/c | AG346840 | 1696 bp | DNA | linear | GSS 02-JUN-2004 |
| LOCUS | | | | | | |
| DEFINITION | | Mus musculus molossinus | DNA, clone:MSMg01-142C12.T7, | | | genomic survey |
| | | sequence. | | | | |
| ACCESSION | | AG346840 | | | | |
| VERSION | | AG346840.1 | GI:47920150 | | | |
| KEYWORDS | | GSS. | | | | |

```

SOURCE
ORGANISM
Mus musculus molossinus
Mus musculus molossinus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
1
AUTHORS
Hattori M., Toyoda, A., Noguchi H., Kojima, T. and Sakaki, Y.
TITLE
BAC end Sequences of Library MSMg01
JOURNAL
Unpublished
AUTHORS
Hattori, M., Toyoda, A., Noguchi, H., Kojima, T. and Sakaki, Y.
TITLE
Direct Submission
JOURNAL
Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail: hattori@gsc.riken.jp, URL: http://hgp.gsc.riken.go.jp/,
Tel: 81-45-503-9114, Fax: 81-45-503-9170)
COMMENT
Clones are derived from the mouse BAC library MSMg01. For BAC
library availability, please contact Kuniya Abe (abe@rtc.riken.jp).
Tsukuba Institute, Bio Resource Center,
The Institute of Physical and Chemical
Research (RIKEN) 3-1-1
Koyadai, Tsukuba, 305-0074 Japan
phone: 81-298-36-9189, fax: 81-298-36-9199
e-mail: abe@rtc.riken.jp
PRIMERS
Sequencing : T7
LIBRARY
Vector : pBACE3.6
R.Site 1 : ECORI
R.Site 2 : ECORI
FEATURES
source
Location/Qualifiers
1. .1696
/organism="Mus musculus molossinus"
/mol_type="genomic DNA"
/sub_species="molossinus"
/db_xref="taxon:57486"
/clone="MSMg01-142C12.T7"
/sex="male"
/tissue_type="mixture of kidney and spleen"
/clone_lib="MSMg01 Mouse Male BAC Library"

```

| ORIGIN | Query Match | 4.3% | Score 88 | DB 9 | Length 1696 |
|--------|-----------------------|--|-------------------|----------------|-------------|
| | Best Local Similarity | 47.3% | Pred. No. 5.4e-06 | | |
| | Matches | 464 | Conservative 0 | Mismatches 501 | Indels 16 |
| Qy | 715 | TAGACTGGTCACTGATAAATAGATATATCTTTAGTAGTAAATATAGTAGGATCTCAATGAC | 774 | | |
| Db | 1290 | TATAATTTAAATAAATAATATCTAAATAAAATAAAATAAATAAATAACATATATATAAAA | 1231 | | |
| Qy | 775 | ATTAATAATTAGAGCTATTTAATTAAAGTTTACTTAATAATAAAGAGAGGTTTAGTAAACAGAAAG | 834 | | |
| Db | 1230 | ATAAAAATACAAAACAACATAATATACAAATAAATAATAAATAATAAATAAATAAATAA | 1171 | | |
| Qy | 835 | CAGGTAAAAACAGAGCTTCCTGCTGTGTGTTTAGTTGTTGTGAGCTCATTTCTTTAAAA | 894 | | |
| Db | 1170 | AATATAATATAAAAAATAAATAAATACTTAACAAAAACAATAAATAACATATACATATA | 1111 | | |
| Qy | 895 | GTAATGTAAACTGATCTFAAGAGCATAGAAATTTAGTACAGGTTAAAACTTTTACAAGAA | 954 | | |
| Db | 1110 | ATATAATATCAATACATCAAAATTTAATAAAAAAAAATAAACAATAAATAAATAA | 1052 | | |
| Qy | 955 | TTTATATTTAAACGAAATATCTTTTATAACATGTCTCTCGGCTGTCAATTAATAAGGATC | 1014 | | |
| Db | 1051 | ATCATATTAATAAATAATCTATCAATAATAAACNATACAAAACAATAATATAAATAATCACC | 992 | | |
| Qy | 1015 | ACTTACTGATCATCCATTTAAAAACCTTTGTTAAAAACAAATTCAGTGAGATAAAATCTTAC | 1074 | | |
| Db | 991 | ATAAATTAATAAATAAATAAATAATAATATAAATAAATAAATAAATAAATAATTT-- | 934 | | |
| Qy | 1075 | AATGAAAAGAGGACAAATGTCTCTTTTGAAAAACAAATAGGTACTCCCTCGCTCCCTCTG | 1134 | | |
| Db | 933 | -----AATAAACAAATATATATAATATATATATAAATAAACAATAACAATATATATA | 880 | | |

```
Qy 1135 AATGTATACATGATTGGACACGGAGACTAAGAAAATGTATTAAGTAATCTAGAGT 1194
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
879 TAATAAATAAATAAATAAATAAATAAACA-TATAAATAATATAAATAAATAACAAA 821
Qy 1195 AAAAAGAAAGAGAAAGTGGTAAAGTAGCGGACCCACCAATATATAATTGATAG 1254
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
820 ATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 761
Qy 1255 ATTTAGAAAAGTAGTTGAAAGTAGTGGTGGTGGGATTTTATATTATAAATAATTTACT 1314
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
760 AAATA-AACATAACAAATAATATAAATAAATAAATAAATAAATAAATAAATAAATA 702
Qy 1315 ATTTTGAGAAAGTTTGAATGTATAGAAATTTGAGTGGGACATCCATAAAGGAAAGTGA 1374
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
701 AAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 642
Qy 1375 TAGAATTAATGGACAGAGGGAGTAACTCTTTATGATATATAAATTTTGTATTG 1434
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
641 TATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 582
Qy 1435 ATTTCAATAGATTATAAATCTATGTTTATAATGATATAATATATTTT-----AAAAATAATAC 1490
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
581 AAAAAACAA-ATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 523
Qy 1491 TATATTAAATCTGATTAGTCGATTACCGCCTTTTATAATTTTACAACTACTGAGTAATG 1550
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
522 AATTATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 463
Qy 1551 AATAAATCAGTTATCTGAAAGCAATAATAATATCTTGTAAACACAGCGTTCGGTCAATGG 1610
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
462 AAGAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 403
Qy 1611 GAAGTTCATGTTATCAATAGTTTAAATAAAGTAAATTTTAAATTAATTTGTTATTT 1670
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
402 AATAAATAAATCAACTAAACAAATAAATAAATAAATAAATAAATAAATAAATAAATA 343
Qy 1671 TTGTTTCAGAAATTTAAATA 1691
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
342 TTATCTCAAAAAATAAAAAAA 322
```

Search completed: March 14, 2005, 20:23:05
Job time : 6856 secs

THIS PAGE BLANK (USPTO)